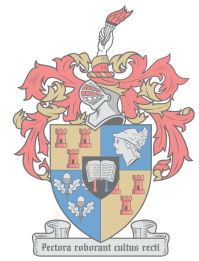
INVESTIGATING THE UTILITY OF LIDAR FOR MODELLING FOREST CANOPY GAPS AND SPECIES CLASSIFICATION

By LEIGHTON LOMBARD (BSC HONS GEOINFORMATICS)

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science at the Stellenbosch University



UNIVERSITEIT
iYUNIVESITHI
STELLENBOSCH
UNIVERSITY

Supervisor: Mr N Poona Co-

Supervisor: Dr R Ismail

March 2018

DECLARATION

By submitting this report electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously, in its entirety or in part, submitted it for obtaining any qualification.

Chapter 3 of this thesis is based on the following publication: Lombard L, Ismail R & Poona N 2017. Modelling forest canopy gaps using LiDAR-derived variables. *Geocarto International* DOI: 10.1080/10106049.2017.1377775.
Leighton Lombard analysed the data and wrote the paper. Dr. Riyad Ismail and Mr. Nitesh Poona contributed to the interpretation of the results and editing of the manuscript.

Date: 13 November 2017

Signature:

SUMMARY

Canopy gaps result from the ineffective use of growing space, presence of features (such as rocks) prohibiting harvesting operations, and naturally by wind, disease, drought, and fires. Traditionally, canopy gaps are detected and interpreted using in situ methods and aerial photography. More recently, remote sensing has been utilized for detecting and delineating canopy gaps. However, optical remote sensing sensors are limited by their spatial resolution. Light detection and ranging (LiDAR) provides new opportunity for canopy gap detection and delineation.

Literature reveals that no study to date has used LiDAR within an object-based image analysis environment (OBIA) to model canopy gaps in South Africa. This research thus aims to investigate the utility of LiDAR for modelling forest canopy gaps and use the delineated canopy gaps for species modelling within a commercial plantation. The first component evaluated the utility of a LiDAR-derived CHM and intensity raster to detect and delineate canopy gaps within a *Eucalyptus grandis* plantation. Canopy gaps were modelled using LiDAR canopy height model (CHM), intensity raster, and a combination of CHM and intensity raster. Thematic accuracies were above 95%, with KHAT values ranging from 0.88 to 0.96. Models were evaluated using an independent test set, yielding thematic accuracies above 90%, with KHAT values ranging from 0.82 to 0.91. A comparative area-based assessment was undertaken on all three datasets and yielded train accuracies ranging from 75% to 95% and test accuracies ranging from 79% to 92%. The combined dataset, i.e. CHM and intensity raster yielded the best overall classification results.

Additionally, delineated canopy gaps were spatially analysed using Getis-Ord Gi* and FRAGSTATS. Getis-Ord Gi* results showed spatial clustering of canopy gaps within the plantation. Furthermore, FRAGSTATS analysed the spatial characterisation of canopy gaps and found varied patch densities (PD) and percentage of landscape (PLAND) occupied by canopy gaps. Canopy gaps were found to be generally irregularly shaped within the plantation. The second component used delineated canopy gaps and LiDAR-derived intensity and texture features to discriminate *Eucalyptus grandis* and *Eucalyptus dunnii* using the random forest (RF) algorithm. Classification models were built using LiDAR intensity and texture information extracted from canopy gaps, and a combination of canopy gaps and forest canopy. Promising results were obtained using a combination of intensity and texture features extracted from canopy gaps alone, with a train out of bag (OOB) error of 7.89 (KHAT = 0.84) and test accuracy of 90.91% (KHAT = 0.81). Improved species discrimination results were obtained using a combination of intensity and texture features and a combination of canopy gaps and forest canopy, with a train OOB error of 3.66 (KHAT = 0.92) and test accuracy of 94.74% (KHAT = 0.88).

The framework developed in this study, i.e. using LiDAR and machine learning, shows promise and robustness, and could potentially assist foresters and forest managers in better understanding the mechanisms underpinning the formation and distribution of canopy gaps. Additionally, this framework shows promise for species discrimination. Therefore, this methodology could potentially be operationalised within commercial forestry for timely and accurate canopy gap detection and species classification.

KEY WORDS

Canopy gaps, LiDAR, OBIA

OPSOMMING

Boomkap gapings word veroorsaak deur die ondoeltreffende gebruik van groeiende ruimte, die teenwoordigheid van kenmerke (soos rotse) wat oesbedrywighede verbied en natuurlik deur wind, siekte, droogte, en brande. Boomkap gapings was tradisioneel opgespoor en geïnterpreteer deur veldwerk en lugfotografie. Onlangs is afstandswaarneming aangewend vir die opsporing en afbakening van boomkap gapings. Optiese sensors vir afstandswaarneming word beperk deur ruimtelike resolusie. Ligopsporing en –verpreiding (LiDAR) bied nuwe geleenthede vir die opsporing en afbakening van boomkap gapings.

Literatuur toon dat geen studie tot dusver LiDAR binne ‘n objekgebaseerde beeldanalise-omgewing (OBIA) gebruik was om boomkap gapings in Suid-Afrika te modelleer nie. Hierdie navorsingsdoel was om LiDAR te ondersoek vir die modellering van boomkap gapings en gebruik die afgebakende boomkap gapings vir spesie modellering binne ‘n kommersiële plantasie. Die eerste komponent het ‘n LiDAR-afgeleide boskap hoogte model (CHM) en intensiteit raster geëvalueer om boomkap gapings in ‘n *Eucalyptus grandis* plantasie op te spoor en te delinieer. Boomkap gapings is gemodelleer deur LiDAR CHM, intensiteit raster en ‘n kombinasie van CHM en intensiteit raster. Tematiese akkuraatheid was bo 95%, met KHAT waardes wat wissel van 0.88 tot 0.96. Modelle is geëvalueer met behulp van ‘n onafhanklike toetsstel, wat tematiese akkuraatheid bo 90% lewer, met KHAT waardes tussen 0.82 en 0.91. ‘n Vergelykende area-gebaseerde assessering is onderneem op al drie datastelle en het oplei akkuraathede opgelewer, wat wissel van 75% tot 95% en toets akkuraathede wat wissel van 79% tot 92%. Die gekombineerde dataset, d.w.s. CHM en intensiteit raster het die beste algehele klassifikasie resultate behaal.

Verder is afgebakende boomkap gapings ruimtelik ontleed met Getis-Ord Gi* en FRAGSTATS. Getis-Ord Gi* resultate het ruimtelike groepering van boomkap gapings in die plantasie getoon. Gevolglik het FRAGSTATS die ruimtelike karakterisering van boomkap gapings ontleed en gevarieerde pleisterdigtheid (PD) en persentasie landskap (PLAND) aangetref. Daar was gevind dat boomkap gapings in die plasie onreëlmatige vorme het. Die tweede komponent gebruik afgebakende boomkap gapings en LiDAR-afgeleide intensiteit en tekstuur-eienskappe om *Eucalyptus grandis* en *Eucalyptus dunnii* te onderskei deur die ewekansige woud (RF) algoritme te gebruik. Sistematiek modelle is gebou met behulp van LiDAR-intensiteit en tekstuur inligting wat uit boomkap gapings uigetrek is en ‘n kombinasie van boomkap gapings en boskap. Belowende resultate is verkry deur die kombinasie van intensiteit en tekstuur eienskappe net uit

boskap gapings te onttrek, met ‘n oplei uit sak (OOB) fout van 7.89 (KHAT = 0.84) en toets akkuraatheid van 90.91% (KHAT = 0.81). Verbeterde spesies diskriminasie resultate is verkry deur ‘n kombinasie van intensiteit en tekstuur informasie en ‘n kombinasie van boskap gapings en boskap met ‘n oplei OOB fout van 3.66 (KHAT = 0.92) en ‘n toets akkuraatheid van 94.74% (KHAT = 0.88).

Die raamwerk wat in hierdie studie ontwikkel is, naamlik die gebruik van LiDAR en masjienleer, toon robuustheid en kan bosbouers en bosbestuurders potensieel help om die vorming en verspreiding van boskap gapings te verstaan. Daarbenewens het hierdie raamwerk belofte vir spesies diskriminasie. Daarom kan hierdie metodologie moontlik binne die kommersiële bosbou aangewend word vir tydige en akkurate boskap gaping en spesies-klassifikasie.

TREFWOORDE

Boomkap gapings, LiDAR, OBIA

ACKNOWLEDGEMENTS

I sincerely thank:

- My family for their continuous support and in particular, my father for assisting with editing and readability of Chapters 3 and 4.
- Ms. Coba Kellerman for her support with editing and contributing to the readability of the thesis.
- The National Research Foundation (NRF) South Africa for funding this project under grant [100802].
- Dr. Rose Masha for doing the final editing of this thesis.
- Finally, my supervisors, Mr. Nitesh Poona and Dr. Riyad Ismail for their invaluable support, assistance, guidance, advice and contribution to the publication of Chapter 3 and the thesis as a whole.

CONTENTS

TABLES

FIGURES

ACRONYMS AND ABBREVIATIONS

CHM	Canopy Height Model
DEM	Digital Elevation Models
DSM	Digital Surface Model
DTM	Digital Terrain Model
GEOBIA	Geographic Image Object-Based Image Analysis
GLCM	Grey-Level Co-Occurrence Matrix
GLDV	Grey-Level Difference Vector
IR	Intensity raster
J-M	Jeffries-Matusita
k-NN	K-Nearest Neighbour
LiDAR	Light Detection And Ranging
LSI	Landscape Shape Index
MRS	Multiresolution image segmentation
OA	Overall accuracy
OBIA	Object-based image analysis
OOB	Out of bag
PD	Patch Density
PLAND	Percentage of Landscape
RF	Random Forest
SEaTH	SEparability and Thresholds
SVM	Support Vector Machines

CHAPTER 1: INTRODUCTION

1.1 COMMERCIAL FORESTRY

Trees are vital resources for our planet and play an important function in ensuring the habitability thereof. Trees and forests can be seen as the lungs of the planet, thus allowing the earth to breathe and maintain oxygen balances. Accordingly, life as we know it would not be possible without the precious natural element (Bredenkamp and Upfold 2012). In South Africa, it is particularly important to preserve and nurture forested areas which endure the semi-arid climate although it is a sparsely treed country (forests utilize only 1.9% of the total land mass) (Bredenkamp and Upfold 2012).

However, the commercial forestry industry provides various job opportunities (employed approximately 75 000 people directly and 500 000 indirectly in 2001) while growing South Africa's foreign exchange markets (Tewari 2001). Economically, the growth of the industry is excellent, and wood demand is expected to increase two-fold. South Africa also has a strategic advantage to other global players in the forestry industry as it is one of the world's leaders in pulp and paper technology (Tewari 2001; Roberts et al. 2007; Bredenkamp and Upfold 2012). In terms of profit, foreign exchange and employment, the South African forestry sector has been especially successful (Tewari 2001). Conversely, the environment has sustained a lot of damage in pursuit of these goals, whereby there is land-degradation, reduction in water resources, loss of biodiversity, a decline of scenic beauty and habitat destruction of many animals (Tewari 2001).

Eucalyptus is grown in commercial plantations in many regions around the world, including the northern and southern hemisphere (Hunter et al. 2004). In South Africa, *Eucalyptus* occupies 48% of commercial hardwoods (DAFF 2012). The fast growth rate and high value timber makes *Eucalyptus* particularly attractive in producing raw material for pulp, paper and other wood products (Sappi 2017). *Eucalyptus* species utilized in commercial forestry include, but are not limited to, *E. grandis*, *E. dunnii*, and *E. smithii*. Both *E. grandis* and *E. smithii* have high growth rates, while *E. dunnii* have a more moderate growth rate. The likelihood of having defects is higher for *E. Smithii*, while *E. dunnii* has less likelihood of defects, and *E. grandis* has the least likelihood of having defects (Sappi 2017).

Commercial forestry within South Africa is mainly found within Mpumalanga and KwaZulu-

Natal provinces and occupies approximately 1% of the national land cover (Peerbhay et al. 2013). Economically and from a forest resource management perspective, it is important to have accurate information relating to the condition and distribution of forested areas (Peerbhay et al. 2013).

1.2 RESEARCH MOTIVATION

Remote sensing has revolutionised forest monitoring and management practices, while providing benefits over aerial image interpretation and in situ field surveys. However, remote sensing should not be seen as a method to replace field surveys; rather, they should be used in a complementary way (Suárez et al. 2005). Remote sensing is building upon and improving traditional methods, whereby satellite imagery provides a synoptic perspective over larger areas with more frequent revisit times. For forest inventory, this technology enables the retrieval of forest attributes of interest at varying accuracies; additionally, methods are becoming more cost effective (Köhl et al. 2006). In recent decades, remote sensing has been effectively used to map forested areas (Ke et al. 2010). Forest species mapping, as well as forest disturbance mapping, have been undertaken using medium spatial resolution sensors (Ke et al. 2010; Malahlela et al. 2014). However, medium spatial resolution sensors are limited to distinguishing forest species at a regional scale. At a finer scale, the potential for inaccuracies with the use of medium resolution sensors increases (Ke et al. 2010). Similarly, for forest disturbance detection, medium spatial resolution sensors are not able to accurately delineate smaller forest gaps (Malahlela et al. 2014).

High spatial resolution sensors are capable of detecting individual trees and smaller forest disturbances (Xie et al. 2008; Ke et al. 2010). These sensors do, however, have limitations i.e. high spatial resolution sensors detect variable spectral reflectance within forest stands comprised of individual tree species (Ke et al. 2010; Malahlela et al. 2014). This results in salt-and-pepper noise in the resulting classification (Ke et al. 2010). Shadows are particularly problematic for forest gap mapping (Hunter et al. 2015).

Light detecting and ranging (LiDAR) overcomes the difficulties faced by passive sensors through multiple laser ranging measurements per m^2 square meter and the ability to generate forest structures both horizontally and vertically (Gaulton and Malthus 2010). Additionally, LiDAR can capture imagery independent of solar illumination (Lillesand et al. 2008). Subsequently, a number of studies have utilized LiDAR for forest species mapping and forest

disturbance detection (Korpela et al. 2010; Gaulton and Malthus 2010; Heinzel and Kock 2011; Li et al. 2013a; Bonnet et al. 2015; Hunter et. al 2015).

1.3 RESEARCH AIM AND OBJECTIVES

The overarching aim of the research was to investigate the utility of LiDAR for modelling forest canopy gaps and use the delineated canopy gaps for species modelling within a commercial plantation.

The specific objectives of the research were to:

1. Detect and delineate forest canopy gaps using a canopy height model (CHM) and intensity raster within an object-based image analysis (OBIA) environment. Additionally, canopy gaps were spatially characterised using Getis-Ord Gi* and FRAGSTATS.
2. Investigate the utility of canopy gap, and LiDAR-derived intensity and texture features, for discriminating *Eucalyptus grandis* and *Eucalyptus dunnii* using the random forest classifier.

1.4 RESEARCH METHODOLOGY

The research is quantitative in nature and utilizes remote sensing, specifically LiDAR for identifying, quantifying, and spatially characterising canopy gaps and investigates whether canopy gaps can be utilized for species classification. Statistical accuracy measures are employed to evaluate the performance of the methodologies presented in chapters 3 and 4. The methodology and flow of the research is presented in Figure 1-1.

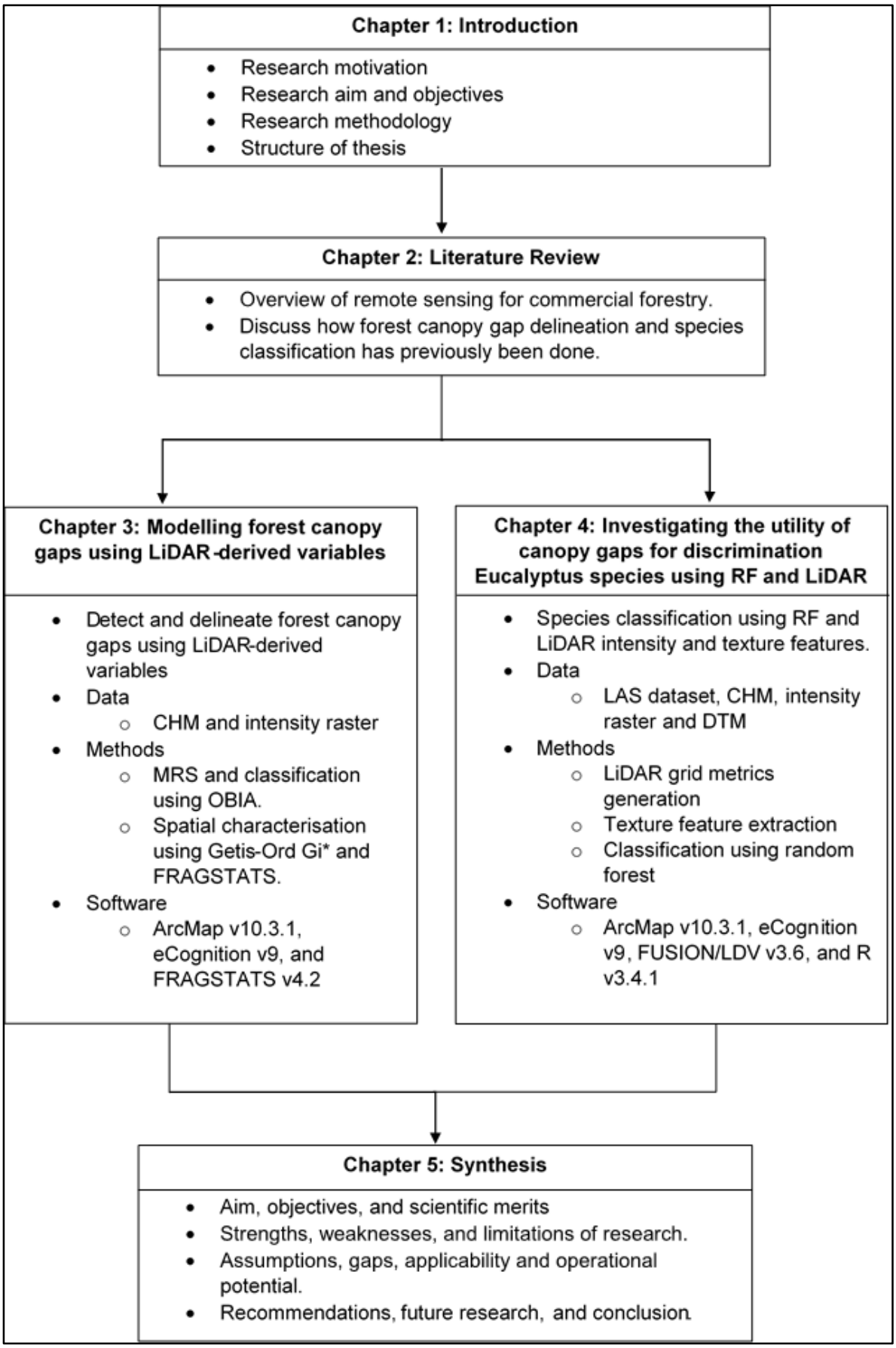


Figure 1-1 Research methodology.

1.5 STRUCTURE OF THESIS

Chapter 1 introduces the research and provides the research motivation, the overarching aim and objectives of the study as well as the research methodology. Chapter 2 contains an in-depth literature review of the relevant concepts and related work. Chapter 3 explains the methodological approach and results of the first research output of the thesis. This chapter specifically focuses on modelling forest canopy gaps using a LiDAR-derived CHM and intensity raster. Further, Getis-Ord Gi* and FRAGSTATS have been used to spatially analyse canopy gaps. Chapter 4 explains the methodological approach and results of the second research output of the thesis. Chapter 4 focuses on species classification using canopy gaps and a random forest classifier. LiDAR intensity and texture metrics have been generated to assist in species discrimination. The final chapter, Chapter 5, is the synthesis chapter. Here the research aim and objectives are revisited; the strengths, weaknesses, and limitations of the research are discussed as well as the operational potential of the techniques. Additionally, recommendations are made for future research and final concluding remarks.

CHAPTER 2: REMOTE SENSING FOR COMMERCIAL FORESTRY

This chapter provides a brief overview of remote sensing and its applicability within commercial forestry. Having knowledge of how remote sensing operates and the varying forms thereof is important in understanding the applicability of remote sensing in different domains. Two main research applications of remote sensing are discussed below, with relevant literature cited. Firstly, a brief overview of remote sensing and remote sensing data processing is provided. Secondly the applicability of remote sensing for canopy gap delineation and forest species mapping is discussed.

2.1 REMOTE SENSING

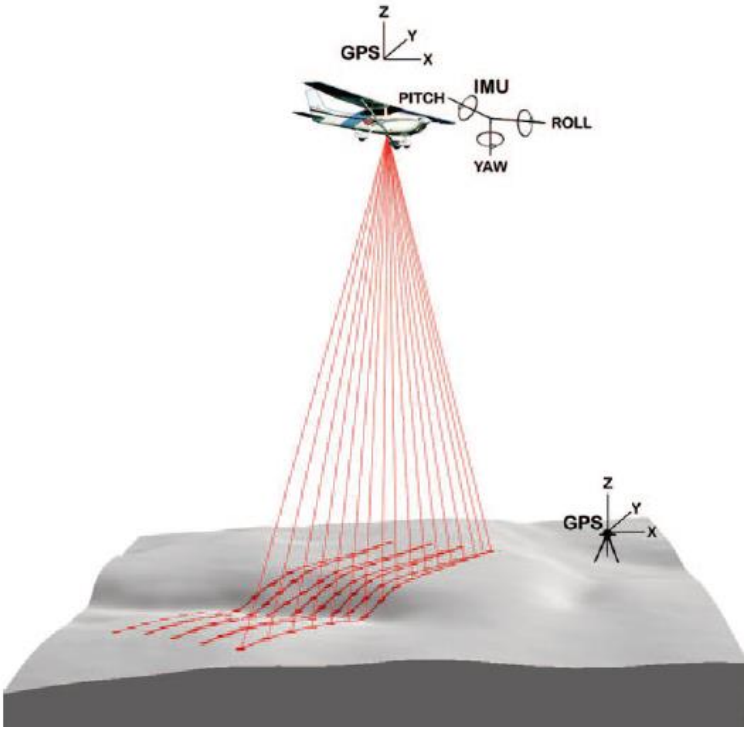
Remote sensing is the practice of obtaining information about the surface of the Earth without being in contact with the surface (Lillesand et al. 2008; Chuvieco and Huete 2010; Jensen 2015). Information can be collected using sensors mounted to an aircraft or satellites (Lillesand et al. 2008; Jensen 2015). Remote sensing includes data acquisition, image processing, image interpretation and subsequent image analysis (Chuvieco and Huete 2010). Information can be captured using a variety of sensors i.e. airborne, ground based, space-borne and radar sensors (Chuvieco and Huete 2010; Jensen 2015). In remote sensing, it is important to understand how electromagnetic energy interacts with a feature on the earth's surface. Having this knowledge enables the extraction of specific information from surface features using remote sensing (Chuvieco and Huete 2010).

Two primary forms of remote sensing sensors exist, these are passive and active sensors. A distinction can be made between passive and active remote sensing systems or sensors. Passive sensors capture information using solar illumination or energy emitted from the Earth's surface (Chuvieco and Huete 2010; Erdle et al. 2011). In contrast, active sensors generate their own energy to acquire information (Erdle et al. 2011). Since passive sensors do not generate their own energy and are dependent on solar illumination, their operational use is limited to the time of day and weather conditions (Fitzgerald 2010). Passive sensors include aerial photography and satellite electro-optical scanners (i.e. cross-track and along-track scanners) (Chuvieco and Huete 2010). Examples of active remote sensing systems include: synthetic aperture radar (SAR), light detection and ranging (LiDAR), sound navigation, and ranging (SONAR) (Baghdadi et al. 2008; Mallet and Bretar 2009; Lillesand et al. 2008; Jensen 2015).

LiDAR sensors use lasers (light amplified by stimulated emission of radiation) to generate a form of light and measure the time it takes the emitted signal to return back to the sensor (Evans et al. 2009; NOAA 2012). LiDAR sensors can be mounted to an aircraft (Figure 2-1), can be satellite-based or, ground-based (Popescu et al. 2011). These sensors are particularly valuable due to their capability to compare characteristics of transmitted energy with the returned signal (Lefsky et al. 2002; Lillesand et al. 2008; Jensen 2015). The timing of pulses, wavelengths, and angles of signals can be assessed. Using this information, a target's structural characteristics can be assessed, but this cannot be achieved using passive sensors (Lillesand et al. 2008).

LiDAR sensors contain three main components: an inertial measurement unit, a global positioning system (GPS) and accuracy clocks (Reutebuch et al. 2005). An inertial measurement unit is used accurately to control aircraft orientation (i.e. roll, pitch, and yaw). A GPS takes accurate readings of a location, and very accurate clocks are used to acquire precise timing of pulses. LiDAR sensors are also capable of generating approximately 150 000 pulses per second, thus resulting in a dense collection of data, which is sometimes referred to as a point cloud (NOAA 2012).

A distinction can be made between small-footprint or large-footprint LiDAR systems (Popescu et al. 2011; Mallet and Bretar 2009). Small-footprint LiDAR illuminates an area of around 0.30m or less, while large-footprint sensors might observe areas of 5m and more (Reutebuch et al. 2005; Hyyppä et al. 2008; Wagner 2010; Mallet and Bretar 2009). The former sensor uses pulsed lasers to generate high-density data (Mallet and Bretar 2009). Large-footprint LiDAR sensors use continuous wave lasers to transmit energy to the Earth's surface and receive a maximum of five returns for each pulse (Chuvieco and Huete 2010).



Source: Reutebuch et al. (2005)

Figure 2-1 Airborne LiDAR.

2.2 REMOTE SENSING DATA PROCESSING

2.2.1 LiDAR processing

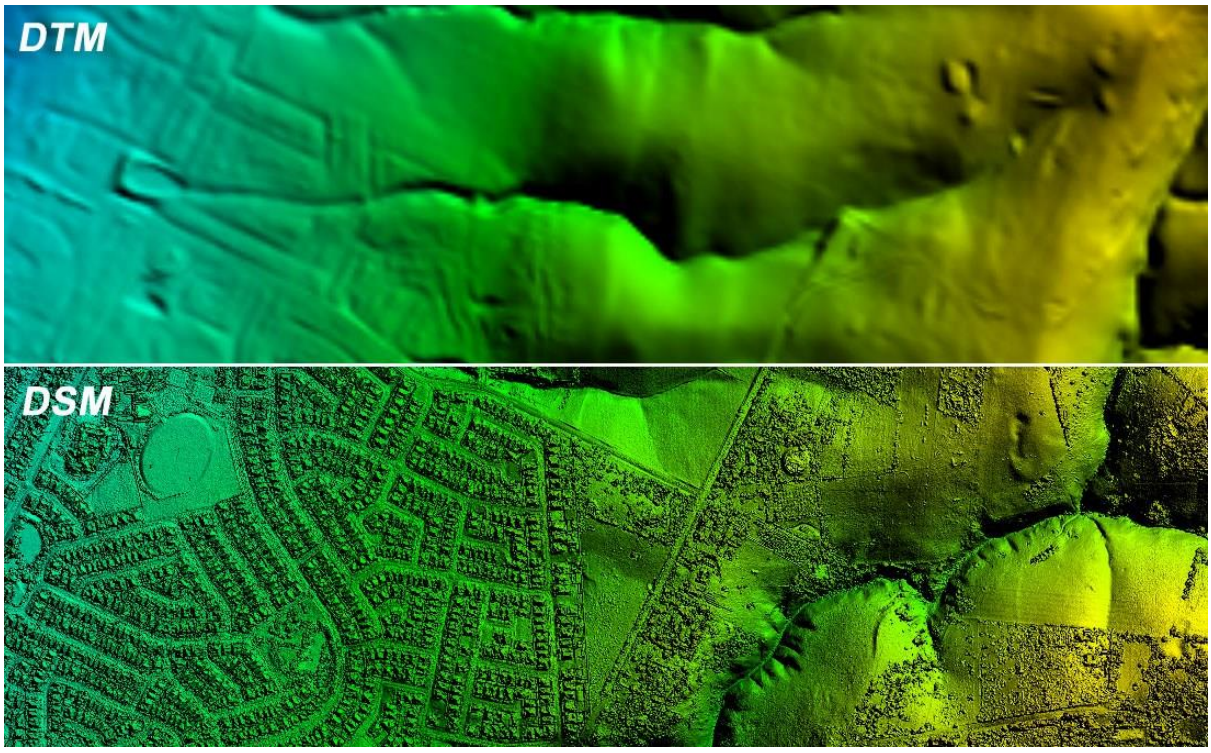
A single transmitted LiDAR signal can return one or multiple return signals (Wagner 2010; Wulder et al. 2012; Jensen 2015). When a transmitted signal interacts with bare ground, one signal is generally returned. However, when a signal interacts with, for example, a tree, multiple backscattered signals will return to the sensor (Jensen 2015). One signal might backscatter from the top branch, another from a lower branch, and a final from the ground next to the tree. In this case, the backscattered signal from the top tree branch will reach the sensor first (known as the 1st return), followed by the signal backscattered from a lower tree branch (i.e. the 2nd return), and lastly, the signal from the ground next to the tree (i.e. the last return). Some sensors are even capable of capturing around five and more returns per emitted signal (Lillesand et al. 2008; NOAA 2012). Furthermore, LiDAR sensors also capture intensity information, where intensity is measured as the return strength of signals as it interacts with surface features (Yunfei et al. 2008; Maltamo et al. 2014).

After LiDAR data has been separated into various returns, the location information validated, and noise removed, it is subsequently stored as X, Y, Z points (Lillesand et al. 2008). Multiple

LiDAR returns can be separated into ground and non-ground returns that can be used to generate digital elevation models (DEM) (Evans et al. 2009; Tinkham et al. 2011). A DEM can further be separated into a digital surface model (DSM) and a digital terrain model (DTM). A DSM represents surface features, whereas a digital terrain model (DTM) represents the bare surface of the Earth (Figure 2-2) (Brovelli et al. 2004; Jensen 2015). Within forested areas, a DTM can be subtracted from a DSM to produce a canopy height model (CHM), where a CHM indicates the height of forest canopy (Räsänen et al. 2014).

A LiDAR generated DSM and DTM have several advantages (Priestnall et al. 2000; St-Onge et al. 2004), namely:

- LiDAR generates highly accurate DSMs and DTMs;
- cost effective elevation model generation; and
- large density of points that enable accurate representation of surface features.



Source: Sing (2013)

Figure 2-2 LIDAR DTM and DSM.

2.2.2 GEOBIA

Prior to the availability of the first software package for geographic image object-based image analysis (GEOBIA) around the year 2000, pixel-based image classification was the predominant remote sensing image classification approach (Blaschke et al. 2014). Pixel-based

image classification is the process of assigning pixels of a feature to a specific class i.e. forests or water bodies (Rahman and Saha 2008; Liu and Xia 2010). GEOBIA (or OBIA) involves aggregating pixels into objects prior to undertaking segmentation and subsequent classification (Navulur 2007; Liu and Xia 2010). An object can be defined as an image entity or scene component that can be distinguished from other entities within an image (Blaschke et al. 2014). Although traditional pixel-based image analysis still provides accurate results, using OBIA features can be discriminated using additional information compared with pixel-based approaches (Navulur 2007; Rahman and Saha 2008; Blaschke 2010; Blaschke et al. 2014):

- A variety of spectral properties can be utilized to characterise an object using OBIA, such as mean and standard deviation; and
- In addition to spectral information, various other properties of objects can be exploited, such as shape (i.e. area, width, or length), texture, and contextual information (i.e. relationship to neighbour objects).

Segmentation is vital step of OBIA and consists of varying methods undertaken prior to classification (Navulur 2007; Rahman and Saha 2008; Drăguţ et al. 2010). Segmentation is the process of aggregating pixels into homogenous regions, called objects, based on shape or spectral information (Drăguţ et al. 2010). There are various segmentation methods available. The popular and widely used method is multiresolution image segmentation (MRS).

Initially MRS considers each pixel to be a unique object. Subsequently, neighbouring objects are merged and form larger objects based on some homogeneity criteria (Rahman and Saha 2008). Objects continue merging to form larger objects until the homogeneity criteria is exceeded (Baatz and Schäpe 2000). A user-defined scale parameter contributes to the homogeneity criteria (Rahman and Saha 2008). The scale parameter determines the size of objects i.e. the larger the scale parameter, the more merging is allowed, thus resulting in larger objects (Drăguţ et al. 2010; Kim et al. 2011b). Furthermore, colour and shape also contribute to the homogeneity criteria. Shape also consists of smoothness and compactness (Rahman and Saha 2008). The larger the influence of colour, the smaller the influence of shape on resulting objects. Similarly, the user can determine whether resulting objects will be more compact or smooth (Drăguţ et al. 2010).

After segmentation, objects can be discriminated using classification by utilizing various object features such as colour, shape, or texture (Gehler and Nowozin 2009). Texture describes the variations of grey tone within objects (Haralick et al. 1973). Within OBIA, texture measures

are largely calculated using a grey-level co-occurrence matrix (GLCM) and a grey-level difference vector (GLDV) (Kim et al. 2011b). The GLCM is a matrix that indicates the probability or number of times combinations of object grey values are found at a specific distance and direction, whereas GLDV are the diagonals of GLCM (Mhangara and Odindi 2013). Various summary statistics can be calculated for both GLCM and GLDV, such as GLCM homogeneity, GLCM mean, GLCM standard deviation, GLDV entropy and GLDV mean (Mhangara and Odindi 2013).

2.2.3 Statistical analysis using Ensemble classifiers

In recent years, there has been an increasing interest in ensemble classifiers (Miao et al. 2012; Rodriguez-Galiano et al. 2012; Belgiu and Drăguţ 2016). Ensemble classifiers use one or more base classifiers to train many more classifiers (Belgiu and Drăguţ 2016). Various classifiers are combined or aggregated using various methods such as voting (Oza and Tumer 2008). Classifiers are trained using bagging, boosting or different variations thereof (Belgiu and Drăguţ 2016). A common distinction between bagging and boosting is that in bagging, a random subset of the total number of training samples is used to train the classifier, whereas boosting uses all samples iteratively to train the classifier (Miao et al. 2012; Belgiu and Drăguţ 2016). Due to the utilization of multiple classifiers in ensembles, an increase in accuracy has been reported compared to more traditional unsupervised and supervised classifiers (Kotsiantis 2011; Rodriguez-Galiano et al. 2012).

A random forest (RF) (Breiman 2001) is an extension of bagging ensemble that combines many tree predictors and randomly selects a subset of trees at each node split (Breiman 2001; Chehata et al. 2009; Miao et al. 2012; Shataee et al. 2012). RF has been reported to obtain improved classification accuracies and is also capable to outperform various classifiers (Liaw and Wiener 2002; Yang et al. 2014; Abdollahnejad et al. 2017; Odindi et al. 2016; Riley et al. 2016). RF includes a set of decision trees based on a bootstrap sample of the data. After selecting the bootstrap sample (referred to as *ntree*), each bootstrap sample grows unpruned classification trees. Thereafter, a random sample of predictors (*mtry*) are chosen at each node. Finally, a majority vote is used and aggregates predictions of *ntree* trees (Liaw and Wiener 2002; Odindi et al. 2016). Several advantages of random forests can be noted such as computationally efficiency of the algorithm as well as being robust against over-fitting and noise (Liaw and Wiener 2002; Poona and Ismail 2014).

Ensemble learning, specifically RF, has been reported in a number of studies using LiDAR data for varying applications such as urban area classification, 3D power-line scene modelling, and landslide detection (Chehata et al. 2009; Kim and Sohn 2010; Chen et al. 2014). The combination of RF and LiDAR has also been utilized in forestry applications (Falkowski et al. 2009; Korpela et al. 2010; Yu et al. 2014; Cao et al. 2016). For example, Falkowski et al. (2009) classified six forest developmental stages using RF in Northern Idaho, USA. Thirty-four LiDAR height metrics were used to distinguish between open stem exclusion, stand initiation, understory initiation, young multistory, mature multistory and old multistory. Using height information, the authors obtained an overall accuracy of 95.54% and KHAT of 93.48.

2.3 FOREST MONITORING USING REMOTE SENSING

Information is a vital component in any decision-making process, and information relating to forestry is no exception (Kangas and Maltamo 2006). Information relating to the extent, quantity and conditions of forests is beneficial to forest planning and policy-making (Reutebuch et al. 2005; Kangas and Maltamo 2006; Köhl al. 2006). Remote sensing is particularly valuable for forestry in this regard (Köhl et al. 2006). Remote sensing is capable of obtaining information about a large area instantly (Köhl et al. 2006). Additionally, remote sensing enables fast retrieval of forest attributes, assists in harvest planning, assesses forest health and forest productivity, open area management, canopy gap delineation, forest species mapping, and decision support applications (Roberts et al. 2007; Bredenkamp and Upfold 2012). However, remote sensing should not be seen as a method to replace field surveys but should be used in a complementary manner (Suárez et al. 2005). Field data are important to validate remote sensing data; similarly, remote sensing data can be used to add value to field surveys (Suárez et al. 2005).

2.3.1 Detecting and delineating canopy gaps using remote sensing

In the past, the main concern of many foresters was timber production above all else, with minimal regard to the environment. This has led to disturbances in forest canopies, resulting in forests with simplified structures, age distribution and species composition (Schliemann and Bockheim 2011). Recently, forestry practices have shifted from a primarily timber production goal to one that has a greater concern for biodiversity preservation and ecosystem functioning. This shift entails implementing cutting practices that mimic natural disturbances (Schliemann and Bockheim 2011; Muscolo et al. 2014; Fox et al. 2000). A number of authors have reported

the benefit of implementing cutting practices that mimic natural disturbances (Fox et al. 2000; Schliemann and Bockheim 2011; Muscolo et al. 2014). This approach produces sufficient harvest yield while restoring forests more naturally (Schliemann and Bockheim 2011). Additionally, canopy gaps have been reported to modify micro-climatic conditions, increase plant diversity and allow regeneration due to light penetrating these openings (Koukoulas and Blackburn 2004; Gaulton and Malthus 2010; Bonnet et al. 2015; Hunter et al. 2015).

Canopy gaps can occur in both natural and commercial forests. The most popular definition of a canopy gap is defined by Brokaw (1982) as a hole in a forested area reaching an average height of about 2m. Canopy gaps can be naturally caused by small-scale disturbances such as the death of one or more trees to larger scale disturbances caused by fires, wind storms, pests, drought or snow (Muscolo et al. 2014). Canopy gaps can also be caused by artificial disturbances, for example, selective harvesting operations (Malahlela et al. 2014). In man-made canopy gaps, the stumps of trees are often still intact, whereas in naturally created canopy gaps, fallen trees may still remain in the gap (Schliemann and Bockheim 2011). Furthermore, canopy gaps can be bare or contain low vegetation (Dietze and Clark 2008) (Figure 2-3).

Figure 2-3 Canopy gap with low vegetation.

Traditionally, canopy gaps were detected and interpreted using ground surveys and aerial photography (Fox et al. 2000). Manual interpretation of canopy gaps involves time consuming and laborious field work (Malahlela et al. 2014; Yang et al. 2015). Ground surveys of canopy

gaps can be done using transects i.e. 25m transects as used by Fox et al. (2000). Using aerial photography has been reported to detect canopy gaps with sufficient accuracy (Fox et al. 2000).

Remote sensing is a valuable alternative to field based canopy gap detection methods and can also detect canopy gaps with improved accuracy (Lippitt et al. 2008; Malahlela et al. 2014; Zielewska-Büttner et al. 2016). Canopy gaps have been delineated with efficient accuracy using both pixel- and object-based approaches with passive sensors (Lippitt et al. 2008; Garbarino et al. 2012; Malahlela et al. 2014; Zielewska-Büttner et al. 2016; Einzmann et al. 2017). Garbarino et al. (2012) used Kompsat-2 to delineate canopy gaps in Lom forest reserve in the Dinaric Alps in Bosnia and Herzegovina. Using a pixel-based approach, a Neural Gas unsupervised classifier and five spectral bands (blue, green, red, NIR, and NDVI), the authors achieved an overall classification of 82%.

Similarly, Zielewska-Büttner et al. (2016) used a pixel-based approach and obtained accuracies as high as 90% and KHAT ranging between 0.66 and 0.88 for canopy gap delineation in the Northern Black Forest, Southwestern Germany. Zielewska-Büttner et al. (2016) defined canopy gaps having a minimum opening of $10m^2$ and maximum height of 2m. Canopy gaps were delineated on a CHM generated from stereo aerial imagery.

Object-based approaches to map canopy gaps have been undertaken by Malahlela et al. (2014) and Einzmann et al. (2017). Malahlela et al. (2014) used WorldView-2 to map canopy gaps in a coastal forest near St. Lucia, KwaZulu-Natal, South Africa. The authors used multiresolution segmentation and six vegetation indices to discriminate between four canopy gap classes (i.e. bare gaps, vegetated gaps, shadow gaps, and others). Indices used include: modified plant senescence reflectance index, normalized pigment chlorophyll index, red edge normalized difference vegetation index, yellow index, near infrared normalized vegetation index and yellow normalized difference vegetation.

Einzmann et al. (2017) detected canopy gaps in two study areas, Munich South and Landsberg, both located near Munich, Germany. The authors used a RapidEye dataset to calculate 175 input features prior to object-based classification. These features include spectral, vegetation, texture and statistical features. The authors used a large-scale mean shift segmentation prior to an RF classifier. Malahlela et al. (2014) and Einzmann et al. (2017) obtained efficient accuracies using OBIA ranging from 93% to 96%. Additionally, Malahlela et al. (2014) compared pixel-based with object-based classification and found the latter to yield more

satisfactory results.

Although high accuracies have been reported with passive remote sensing sensors, various limitations have been reported. Some of these include saturation of the visible-near infrared signal in dense vegetation as well as the presence of shadows (Malahlela et al. 2014; Hunter et al. 2015; Yang et al. 2015). LiDAR sensors can be used to overcome these limitations by generating accurate height information (Koukoulas and Blackburn 2004; Malahlela et al. 2014; Hunter et al. 2015).

Gaulton and Malthus (2010) compared pixel-based canopy gap delineation using a LiDAR point cloud method with a LiDAR-derived CHM. The authors reported an increase in accuracy using the point cloud approach with an overall accuracy of 78.20% compared to an 74.50% overall accuracy obtained using a CHM. Despite the increase in accuracy of the point cloud approach, canopy gap delineation using the CHM was found to be less computationally intensive (Gaulton and Malthus 2010; Bonnet et al. 2015). Use of a LiDAR-derived CHM in a OBIA environment for canopy gap delineation was tested by Vepakomma et al. (2008) and Bonnet et al. (2015).

Using a region grow segmentation algorithm, Vepakomma et al. (2008) delineated canopy gaps in Lake Duparquet Teaching and Research Forest, Canada. Bonnet et al. (2015) used the following criteria for canopy gap delineation in Wallonia, Belgium: a canopy gap must have a minimum area of $50m^2$, minimum width of 2m, and maximum height of 3m. Using this criteria, Bonnet et al. (2015) tested three mapping methods i.e. threshold, pixel-based, and object-based supervised classification using multiresolution classification. Both Vepakomma et al. (2008) and Bonnet et al. (2015) obtained sufficient accuracies for object-based canopy gap delineation, with overall accuracies above 79%.

According to literature, LiDAR is effective for delineating and mapping canopy gaps, while avoiding the effects of shadows and saturation of signal in dense vegetation. Using a combination of LiDAR and OBIA has been reported to obtain improved canopy gap mapping accuracies (Vepakomma et al. 2008; Bonnet et al. 2015). However, limited literature was found on canopy gap mapping in South Africa. Subsequently, Chapter 3 documents canopy gap mapping using both LiDAR and OBIA in a commercial forest in KwaZulu-Natal, South Africa.

2.3.2 Forest species discrimination using remote sensing

The ability to model forest species are important for a number of reasons. These include identifying the impacts of forest disease and pathogens, ensuring the optimal growth and productivity of commercial species by matching the traits of species to specific climatic and environment conditions, disturbance detection, as well as creating species-specific growth and yield models or treatment schedules (Wilson et al. 2012; Peerbhay et al. 2013; Maltamo et al. 2014). Additionally, forest species discrimination is important for mapping species diversity in habitat modelling and management decision-making (Immitzer et al. 2012).

Field observations, periodic surveys and aerial photography are traditional approaches to acquire forest species information (Peerbhay et al. 2013). Forest stands were manually delineated using aerial photographs (Van Coillie et al. 2007; Kim et al. 2009a). This is a laborious process of classifying forest types (Kim et al. 2009a). In addition to the laborious manual process of delineating forest stands, traditional field surveys and aerial observation techniques are also substantially time consuming and costly (Van Coillie et al. 2007; Immitzer et al. 2012; Peerbhay et al. 2013). Despite these obstacles, a number of studies have reported satisfactory results using aerial photography for forest species mapping (Aldred and Hall 1975; Meyer et al. 1996; Key et al. 2001; Olofsson et al. 2006).

As early as in 1975, Aldred and Hall (1975) achieved efficient accuracies for delineating balsam fire, white spruce, red pine, white pine, and jack pine tree species using large-scale photos in western Quebec, north of Maniwaki. The authors used two main methods for species discrimination. Firstly, forest plot boundaries were delineated. Thereafter, random plots were selected while trees above 10m were manually numbered and classified based on specific species attributes. The authors reported accuracies of approximately 90%.

Meyer et al. (1996) utilized colour infrared-aerial imagery to map Spruce, Pines, Fir, and Beech in Canton of Aargau, Switzerland. The authors utilized two classifiers i.e. maximum likelihood and parallelepiped classifiers on two datasets. This was undertaken on the original three spectral bands as well as on a modified six image band dataset. Features or bands in the latter dataset included: the original three image bands (i.e. green, red, and infrared), brightness, texture, and a 5x5 centre-weighted low pass filter feature. Tree species classification using parallelepiped image classification procedure reported an improved accuracy compared with maximum likelihood with an overall accuracy of 87%. Olofsson et al. (2006) also used aerial

imagery for species discrimination. Olofsson et al. (2006) obtained a slight improvement in accuracy compared with the results of Meyer et al. (1996). The authors reported an overall accuracy of 89% using discriminant analysis to model scots pine, Norway spruce and deciduous trees in Remningstorp estate in south-western Sweden.

In addition to spectral information of aerial photographs, Key et al. (2001) utilized a multi-temporal dataset to discriminate four deciduous tree species in West Virginia, USA. These species include: yellow poplar, white oak, red oak, and red maple. By using multi-temporal data, a large number of features were included in the classification. An overall classification accuracy of 74% was obtained using the maximum likelihood classification. Despite the increased number of features, Key et al. (2001), Meyer et al. (1996), and Olofsson et al. (2006) obtained improved results using a smaller number of features.

The advent of advanced satellite and airborne sensors has improved the efficiency of forest species discrimination (Immitzer et al. 2012; Maltamo et al. 2014; Wang et al. 2016). Using the RF classifier to delineate ten tree species in Austria was undertaken by Immitzer et al. (2012). An overall accuracy of 82% was obtained using WorldView-2 to discriminate forest species using spectral information. In KwaZulu-Natal, South Africa, Odindi et al. (2016) utilized RapidEye imagery in combination with a RF classifier for alien and indigenous vegetation mapping. Dominant species were discriminated, with accuracies ranging between 68% and 80%. Yang et al. (2014) also used RapidEye imagery for species mapping by utilizing RF and support vector machines (SVM). The authors reported that RF outperformed SVM. More recently, Abdollahnejad et al. (2017) used Quickbird imagery to delineate forest species using a combination of classifiers. Similar to the findings of Yang et al. (2014), Abdollahnejad et al. (2017) also reported RF as having outperformed SVM and k-nearest neighbour for species discrimination, with an overall accuracy of 63.85%.

According to Wolter and Townsend (2011), pixels have difficulty discriminating forest species, particularly for high spatial resolution passive sensors. The advent of airborne LiDAR has been described as a valuable addition for forest species discrimination (Ørka et al. 2009; Korpela et al. 2010; Maltamo et al. 2014). More specifically, LiDAR-derived height and intensity information can be beneficial for species classification (Ørka et al. 2009). Subsequently, several studies have explored LiDAR for forest species mapping. See for example, Donoghue et al. (2007), Reitberger et al. (2008), Säynäjoki et al. (2008), Wagner et al. (2008), Korpela et al. (2010), van Leeuwen and Nieuwenhuis (2010), Zhao et al. (2011), Simonson et al. (2012),

Li et al. (2013a), and Maltamo et al. (2014).

For example, Donoghue et al. (2007) used small footprint LiDAR intensity and height metrics to discriminate spruce and pine. The authors utilized varying LiDAR metrics, including: mean height, coefficient of variation, skewness, and percentage of ground returns. Reitberger et al. (2008) obtained overall accuracies as high as 96% using small-footprint full waveform LiDAR for discriminating coniferous and deciduous species using unsupervised classification. Säynäjoki et al. (2008) differentiated European aspen from other deciduous trees with an accuracy of 78.6%. The authors calculated height percentiles, proportion of laser points at each percentile, proportion of vegetation hits, mean height, standard deviation of height, mean intensity, standard deviation of intensity and intensity percentiles.

A vast amount of literature can be found on species classification using LiDAR-derived information extracted from forest canopy. In contrast, only one study to date has utilized gap information in combination with forest canopy information to assist in species classification (Li et al. 2013a). For example, Li et al. (2013a) used LiDAR-derived metrics from both forest tree crowns and intra-tree crown gaps to assist in species classification in Ontario, Canada. The authors utilized a gap distribution metric as well as 3-D texture, relative foliage degree and relative scale of foliage clustering extracted from forest canopy. Chapter 4 of this research aims to contribute to this gap in the literature by utilizing both canopy gap and forest canopy information for species discrimination. Additionally, this chapter also utilizes the efficiency of the RF algorithm and LiDAR data for species modelling, as shown by a number of studies (Korpela et al. 2010; Yu et al. 2014; Cao et al. 2016).

CHAPTER 3: MODELLING FOREST CANOPY GAPS USING LIDAR-DERIVED VARIABLES

3.1 ABSTRACT

Remote sensing has revolutionised forest management and has been widely employed to model canopy gaps. In this study, a canopy height model (CHM) and an intensity raster (IR) derived from light detection and ranging (LiDAR) data were used to model canopy gaps within a four-year-old *Eucalyptus grandis* forest using an object-based image analysis (OBIA) approach. Model thematic accuracies using the CHM, intensity raster, and combined dataset (CHM and IR) were all above 90%, with KHAT values ranging from 0.88 to 0.96. Independent test thematic accuracies were also above 90%, with KHAT values ranging from 0.82 to 0.91. A comparative area-based assessment yielded accuracies ranging from 70% to 90%, with the highest accuracies achieved using the combined dataset. The results of this study show that using a CHM and intensity raster, and an OBIA approach, provides a viable framework to accurately detect and delineate canopy gaps within a commercial forest environment.

3.2 INTRODUCTION

Canopy gaps in plantations occur due to the ineffective use of growing space, the presence of features (such as rocks) that prohibit planting, and harvesting or thinning operations (Vehmas et al. 2011; Malahlela et al. 2014; Bonnet et al. 2015). Additionally, the impact of natural disturbances by wind, snowfall, disease, drought, climate change, and fires lead to the formation of canopy gaps (Vehmas et al. 2011; Muscolo et al. 2014; Bonnet et al. 2015). Tree mortality is usually higher in young stands with high stocking levels. A tree thus has a higher probability of dying in a denser stand, in particular on poor quality sites (Mabvurira and Miina 2002).

Canopy gaps affect a forest in various ways. Canopy gaps lead to variations in light conditions, temperature, soil moisture, and nutrient availability (Negrón-Juárez et al. 2011; Muscolo et al. 2014). Additionally, canopy gaps promote both biodiversity - and pedodiversity and have an important influence on forest dynamics (Garbarino et al. 2012). For example, canopy gaps increase light penetration to the understory, providing an opportunity for enhanced growth of the canopy (Garbarino et al. 2012; Gray et al. 2012), as well as stimulating the growth and survival of native species (Muscolo et al. 2014).

Canopy gaps are, however, difficult to detect (Vepakomma et al. 2008) as they are often varied in size, and generally dominated by low vegetation, such as sprouts (Dietze and Clark 2008). Several studies have explored the utility of multispectral remote sensing imagery to detect and quantify canopy gaps. For example, Lippitt et al. (2008) mapped selective logging / harvesting, which lead to canopy gaps (Vajari et al. 2012), using multi-temporal Landsat Enhance Thematic Mapper Plus imagery. Of the five machine learning algorithms tested, classification trees yielded the highest overall accuracy (94%) and KHAT value (0.60). Garbarino et al. (2012) successfully used high-resolution Kompsat-2 imagery to detect canopy gaps in the Dinaric Alps in Bosnia and Herzegovina. The authors applied an unsupervised artificial neural network, coupled with field measurements to detect 650 canopy gaps with an overall accuracy of 82%. More recently, Zielewska-Büttner et al. (2016) used a canopy height model (CHM) derived from multi-temporal stereo aerial imagery to detect canopy gaps in the Northern Black Forest, Southwest Germany. The authors achieved overall accuracies ranging from 82% to 90% as well as KHAT values ranging from 0.66 to 0.88.

The utility of passive remote sensing systems is, however, limited by its spatial resolution. The medium to coarse spatial resolution makes discriminating medium to small sized canopy gaps, difficult (Malahlela et al. 2014). The use of high spatial resolution multispectral sensors, e.g. QuickBird can help alleviate this problem. However, these higher resolution sensors tend to suffer from saturation of the visible-near infrared signal in dense vegetation (Malahlela et al. 2014). Additionally, shadows provide an added setback, especially for high-resolution sensors (Hunter et al. 2015). Espírito-Santo et al. (2014) noted that shadows are a significant problem in tropical areas, where shadows occur in both gap and non-gap areas. The authors concluded that detecting and mapping canopy gaps using spectral information is challenging.

The advent of active remote sensing systems such as LiDAR (light detection and ranging) provides an alternative to the passive remote sensing systems. More importantly, LiDAR overcomes many of the obstacles (saturation and spatial resolution) faced by passive remote sensing systems (Frolking et al. 2009; Malahlela et al. 2014; Hunter et al. 2015). A key advantage of LiDAR is the provision of forest structural information in both the horizontal and vertical domain (Gaulton and Malthus 2010). Having both horizontal and vertical information, allows for detecting canopy gaps beneath the outer circumference of tree branches (Vehmas et al. 2011).

Several studies have successfully used LiDAR for modelling canopy gaps. For example,

Gaulton and Malthus (2010) used two approaches, i.e. a LiDAR point cloud with local maxima filtering and clustering, and a LiDAR-derived CHM to identify canopy gaps in three *Picea sitchensis* plantations in Scotland. The mean overall classification accuracy using the CHM was 75%, compared with 78% using the LiDAR point cloud. Additionally, the CHM produced a higher error (RMSE = 38%) in identifying total canopy gap area, compared with the LiDAR point cloud (RMSE = 22%). However, the authors concluded that using the LiDAR point cloud was more computationally intensive. More recently, Hunter et. al (2015) employed multi-temporal LiDAR to analyse temporal changes in canopy gap size and frequency within two sites in Tapajos National Forest, Brazil.

More recently, several authors (for example Vepakomma et al. 2008; Malahlela et al. 2014; Bonnet et al. 2015; Einzmann et al. 2017) have employed object-based image analysis (OBIA) for detecting and delineating canopy gaps. OBIA involves analysing images using objects instead of pixels. An image is first segmented into multiple objects, prior to image classification (Navulur 2007). The key advantage to using objects for classification is the additional attribute information such as shape, texture, and morphology (Navulur 2007). The use of additional object attributes have shown to increase classification accuracy (Malahlela et al. 2014).

Malahlela et al. (2014) mapped canopy gaps in an indigenous subtropical coast forest using high-resolution WorldView-2 imagery. The authors tested both a pixel-based and object-based classification approach. The object-based approach yielded an overall accuracy of 94% compared with the pixel-based approach, which yielded an accuracy of 87%. Einzmann et al. (2017) successfully used an object-based approach to map canopy gaps caused by windthrow, with high-resolution RapidEye imagery. The authors achieved overall accuracies ranging from 93% to 96%. Using a LiDAR-derived CHM, Vepakomma et al. (2008) successfully mapped canopy gaps using OBIA. The authors achieved an overall accuracy of 96%. More recently, Bonnet et al. (2015) used a LiDAR-derived CHM, slope of the CHM, and a canopy porosity index to map canopy gaps in the watershed of the Houille River, Belgium. The authors tested three mapping methods, i.e. thresholding, supervised classification, and per-object supervised classification yielding overall accuracies ranging from 62% to 82%.

A review of the literature showed that no study to date has used LiDAR derivatives, i.e. CHM and intensity raster, within an OBIA environment, to model canopy gaps in a commercial plantation in South Africa. It is within this context that this study aims to evaluate the utility of a CHM and an intensity raster to detect and delineate canopy gaps within a *Eucalyptus grandis*

forest. Additionally, we employ spatial statistics to analyse and characterise the spatial nature of canopy gaps within our study area. These analyses can assist foresters and forest managers in better understanding the mechanisms underpinning the formation and distribution of canopy gaps, and form part of an integrated forest management framework.

3.3 MATERIALS AND METHODS

3.3.1 Study Area

The Sappi Riverdale plantation (Figure 3-1) is located near the town of Richmond in the KwaZulu-Natal midlands, South Africa. The region has a mild, warm and pleasant climate, with a mean annual temperature of 17.4°C and mean annual precipitation of 872mm. Riverdale is approximately 5999ha of *Eucalyptus* forest. A total of 27 compartments are within the plantation, comprised of *E. grandis* (*n* = 15), *E. smithii* (*n* = 2), and *E. dunnii* (*n* = 10) (Macfarlane 2006). For this study we focus on *E. grandis* contained in Block E and Block F.

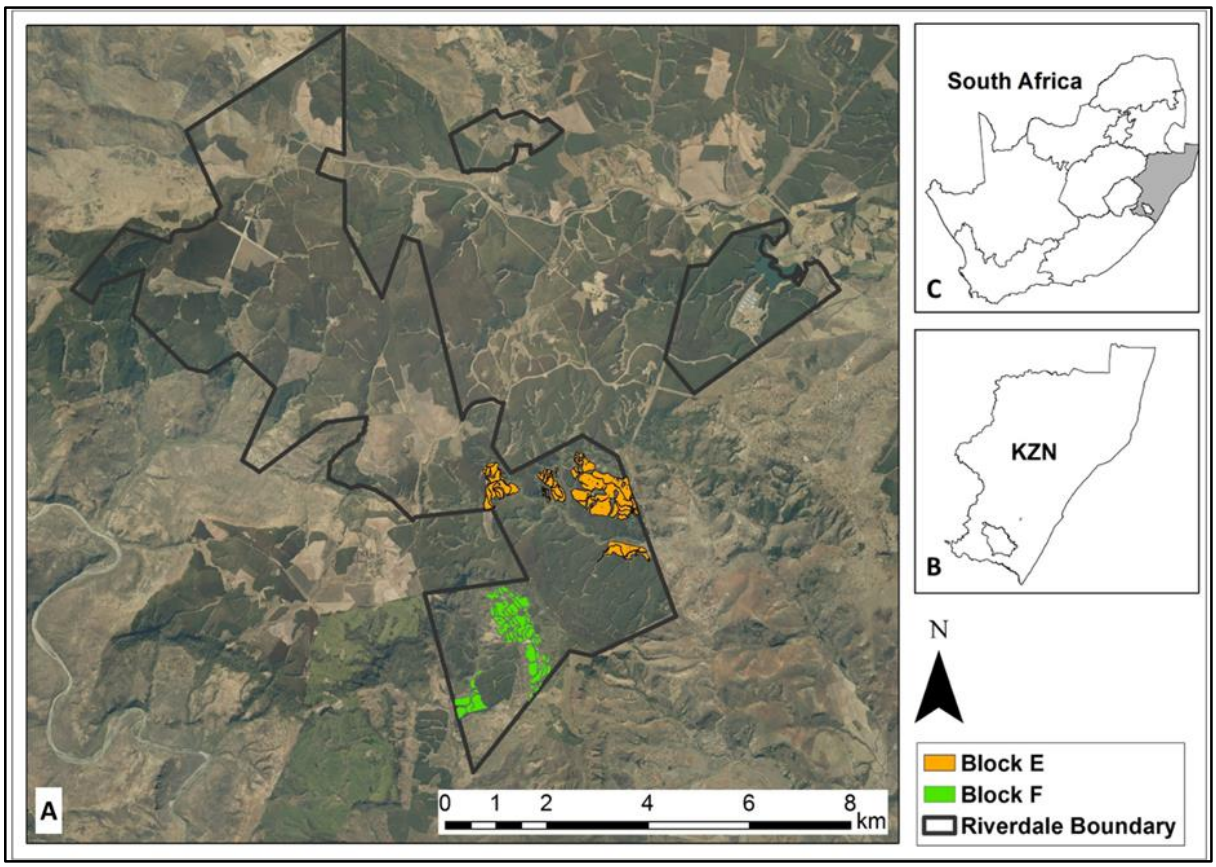


Figure 3-1 The Riverdale plantation (a) located near Richmond in the province of KwaZulu-Natal (b), South Africa (c). Background image is ESRI ArcGIS online's 50 cm colour imagery for South Africa.

3.3.2 Image and field data

A LiDAR survey using a Leica ALS50-2 laser scanner with multi-pulse was conducted from the 15th to the 22nd of March 2014. The number of returns was four at a pulse rate of 126 000Hz and scan rate of 53Hz. The average flying height was 820m with a 50% minimum flight line overlap. The LiDAR data was used to create an intensity raster using the LAS Dataset to Raster tool in ArcMap v10.3.1 (ESRI 2015). A low pass filter with a 3x3 kernel was used to remove the inherent noise in the intensity raster (ESRI 2015). Aerial imagery of 15cm spatial resolution was also acquired on 12 April 2014 at an average flying height of 213m. The aerial imagery served as reference data for undertaking the accuracy assessments. Field data in the form of enumerated plot data for each compartment were provided by Sappi Forests. The blocks and compartments used for the analysis was selected using the field data. Compartment F1 was used for training, i.e. model building, whereas compartment F3a was used as an independent test dataset. A 1m resolution CHM was also provided by Sappi Forests.

3.3.3 Canopy gap delineation using multiresolution segmentation (MRS)

Multiresolution segmentation (MRS) was utilized for delineating canopy gaps using the CHM, intensity, and a combined CHM and intensity raster in eCognition developer 9 (Definiens 2007). MRS is a region-merging segmentation approach that forms larger homogenous image objects by iteratively merging singular image objects (Definiens 2007; Varo-Martínez et al. 2017). MRS starts with image pixels that are aggregated into image objects (Baatz and Schäpe 2000). Region-merging is achieved using a homogeneity criterion that measures how similar or dissimilar an image object is. The homogeneity criterion is influenced by the colour and shape of image objects, with values ranging between 0 and 1 (Baatz and Schäpe 2000; Definiens 2007). Colour describes the influence of spectral values, whereas shape determines the level of smoothness and compactness of image objects (Navulur 2007). To determine the optimal scale, shape, and compactness parameter values, we tested scale factors of 20, 10 and 5, coupled with shape and compactness values ranging from 0.1 to 0.9, in increments of 0.1. Combinations of scale, shape, and compactness parameter values were tested using the CHM, intensity raster, and combined dataset.

3.3.4 Rule-based classification

To determine the optimal thresholds for OBIA, we employed the SEparability and THresholds (SEaTH) (Nussbaum et al. 2006) tool. SEaTH employs the Jeffries-Matusita (JM) distance

(Equation 3-1 and Equation 3-2), which measures the separability (ranging from zero to two) between two classes, i.e. forest and canopy gaps based on training samples. For normally distributed classes, this JM distance is stated as (Richards and Jia 1999):

$$JM_{ij} = 2(1 - e^{-B}) \qquad \text{Equation 3-1}$$

in which

$$B = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{C_i + C_j}{2}\right)^{-1} (\mu_i - \mu_j) + \frac{1}{2}\ln\left(\frac{(|C_i + C_j/2|)}{\sqrt{|C_i| * |C_j|}}\right) \qquad \text{Equation 3-2}$$

where

B = Bhattacharyya distance

i and j = two classes being compared (i.e. forest and canopy gaps)

C_i = the covariance matrix of signature i

μ_i = the mean vector of signature i

ln = the natural logarithm function

$|C_i|$ = the determinant of C_i

The closer the JM value is to 2, the better the separability between forest and canopy gaps, whereas a value lower than 1.8 indicates that the two classes are less separable. Additionally, SEaTH avoids the time constraining trial-and-error process of manually testing thresholds for ruleset development (Gao et al. 2011). SEaTH generates thresholds for classification by using object statistics derived from a representative selection of training areas per class (Nussbaum et al. 2006). The resulting threshold values are produced using a Gaussian probability mixture model (Nussbaum et al. 2006).

The SEaTH tool has previously been used by Gao et al. (2011) for rule-based land cover mapping with Landsat-8 Enhance Thematic Mapper Plus imagery and Huang et al. (2015) for rule-based classification of forest stands with QuickBird imagery. We used SEaTH to statistically identify optimal thresholds for classifying canopy gaps. A representative training sample from compartment F1 (n = 6) and compartment F3a (n = 5) was used to calculate object statistics used in SEaTH. The resulting SEaTH thresholds were used to build classification rules in eCognition for classification of the CHM, intensity raster, and combined dataset.

3.3.5 Accuracy assessment

We evaluated both the thematic accuracy and comparative area accuracy of the delineated canopy gaps. To evaluate the thematic accuracy of the delineated canopy gaps, we used 80 reference points and a confusion matrix. The 80 reference points were generated using the create random points tool in ArcMap (ESRI 2015). A confusion matrix provides several measures of accuracy including errors of commission, errors of omission, and overall accuracy. Additionally, a multivariate statistic called KHAT, was used to test agreement between the test and training data (Congalton and Green 2009).

To determine the percentage match between the reference canopy gap area (digitised from the very high resolution aerial photographs) and the OBIA delineated canopy gap area, a comparative area-based assessment (Champion et al. 2008; Hermosilla et al. 2011; Gomes and Maillard 2013) was employed. Reference canopy gap areas were manually digitised using aerial imagery. Overall accuracy for the comparative area-based assessment (Equation 3-3) was calculated as the percentage of total delineated canopy gap area relative to the total reference canopy gap area:

$$\frac{total\ delineated}{total\ reference} * 100 \qquad \text{Equation 3-3}$$

3.3.6 Spatial statistics

3.3.6.1 Assessing spatial clustering using Getis-Ord Gi*

Hotspot analysis (Swetnam et al. 2015; Reddy et al. 2016) was used to test for spatial clustering of canopy gaps. A hotspot analysis uses the Getis-Ord Gi* to measure spatial clustering of a sample and how it varies from an expected value (Getis and Ord 1992; Reddy et al. 2016). The Getis-Ord Gi* statistic calculates z-scores and p-values that indicate spatial clustering of high data values (i.e. high z-score and low p-value) and low data values (i.e. low negative z-score and low p-value) (ESRI 2015; Swetnam et al. 2015; Reddy et al. 2016). Within the context of this study, we are interested in clustering of large canopy gaps (i.e. locations having a high z-score and low p-value) within blocks E, F, and the combined block (E + F). A p-value greater than or equal to 0.1 was considered insignificant and therefore, not a hotspot.

The Getis-Ord Gi*statistic (Equation 3-4) is calculated as follows (Reddy et al. 2016):

$$G_I^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \overline{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}}$$ Equation 3-4

Where: $\overline{X} = \frac{\sum_{j=1}^{n} x_j}{n}$ and $S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - (\overline{X})^2}$; x_j represents the attribute value for j, $w_{i,j}$ = the spatial weight between features i and j, and n = the number of features.

The Hotspot analysis tool in ArcMap (ESRI 2015) requires a distance band, which represents the scale of analysis. The distance band was determined using incremental spatial autocorrelation (ESRI 2015) for block E, block F, and the combined block.

3.3.6.2 Spatial characterisation of canopy gaps using FRAGSTATS

Four spatial statistical metrics were computed (Table 3-1) on the class level for block E, block F, and the combined block. Class level is equivalent to the all patches of a certain class (i.e. canopy gaps) within a block. A neighbour rule can be specified for delineating patches (McGarigal et al. 2012). Patch membership can be assigned using either a 4-cell rule or 8-cell neighbour rule. We used the 8-cell rule, where eight adjacent cells are considered, including 4 orthogonal and 4 diagonal neighbours (McGarigal et al. 2012). The metrics were calculated using the FRAGSTATS package, version 4.2 (McGarigal et al. 2012).

Table 3-1 Spatial metrics used computed at the class (block) level (adapted from McGarigal et al. (2012)).

Metric	**Description**	**Equation**
Percentage of Landscape (PLAND)	PLAND measures the percentage of canopy gap within a block.	$P_i = \frac{\sum_{j=1}^{n} a_{ij}}{A}(100)$ P_i = percentage of block occupied by canopy gap *i*. a_{ij} = area of canopy gap *ij* (m^2). A = total area of the block (m^2).
Shape Index	Shape index determines the size of a canopy gap on patch level, and mean size of all canopy gaps on block level or class level. For a square canopy gap, the shape index equals 1; the more irregular the shape, the larger the shape index.	$SHAPE = \frac{0.25\, P_{ij}}{\sqrt{a_{ij}}}$ P_{ij} = perimeter of canopy gap *ij* (m). a_{ij} = area of canopy gap *ij* (m^2).
Patch Density (PD) (per 100 hectares)	PD refers to the number of canopy gaps within a class or block.	$PD = \frac{n_i}{A}(10000)(100)$ n_i = number of canopy gaps in the block for class *i*.
Landscape Shape Index (LSI)	LSI is similar to shape index. However, LSI treats canopy gaps as a single entity. LSI computes the perimeter-to-area ratio for a given entity. LSI equals 1 when the entity comprises a singular square shape, and increases with increasing irregularity.	$LSI = \frac{0.25E^*}{\sqrt{A}}$ E* = total length (m) of edge in block; includes the entire block boundary and some or all background edge segments. A = total block area (m^2).

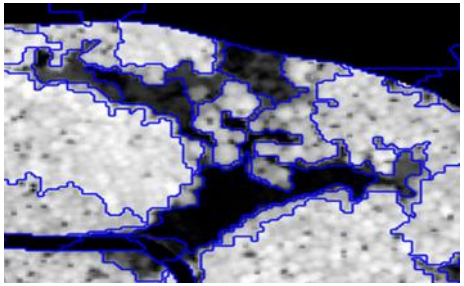
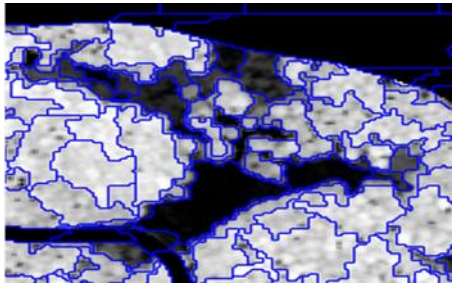
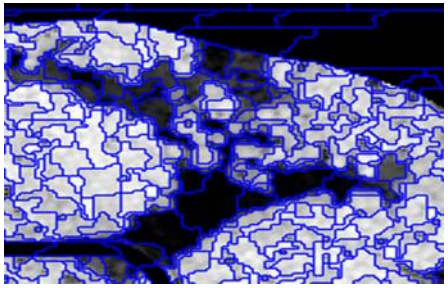
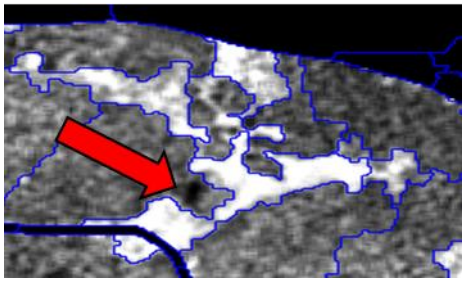
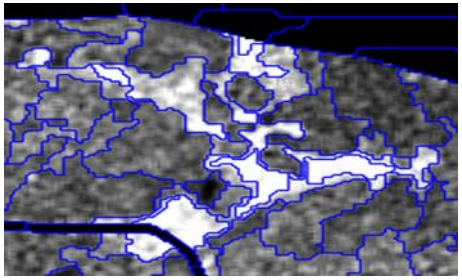
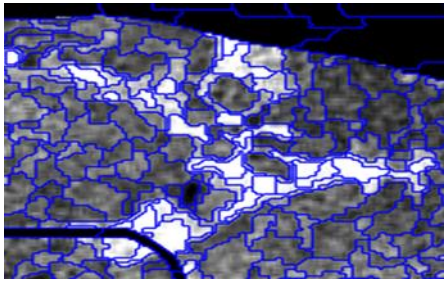
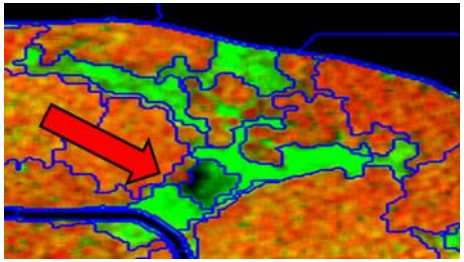
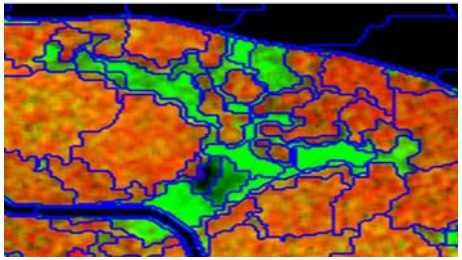
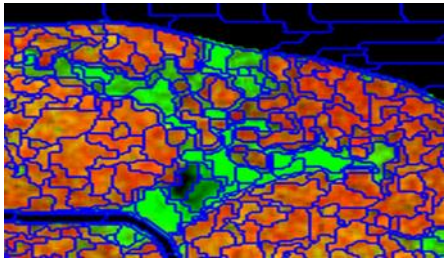
3.4 RESULTS

3.4.1 Canopy gap delineation and classification

The optimal shape and compactness parameter values for delineating both small and large canopy gaps were 0.1 and 0.5 respectively. Table 3-2 shows the MRS results using varying scale factors with a shape parameter value of 0.1 and compactness parameter value of 0.5. It is evident that at a given scale, the resulting image objects were similar in size for all three datasets. For example, at a scale of 10, the resulting image objects for the CHM, intensity raster, and combined dataset are similar in size and number. However, as the scale factor decreased, the size of the resulting image objects also decreased. For example, the size of resulting image objects produced for the combined dataset at a scale of 5 were smaller, compared with the size of resulting image objects produced for the combined dataset at a scale of 20. Additionally, a decrease in the size of image objects resulted in an increase in the number of image objects.

An inherent limitation of using only the CHM is evident from its inability to identify and differentiate water bodies from canopy gaps. However, the water body (indicated by the red arrow) in Table 3-2 was identified and differentiated from a canopy gap using the intensity raster. The ability of the intensity raster to identify and differentiate the water body was attributed to the low intensity of the water body compared with the surrounding area, i.e. forest.

Table 3-2 MRS of CHM, intensity, and combined dataset at scale factors 20, 10 and 5.

Datasets/ Scale factors	20	10	5
CHM			
Intensity			
CHM + Intensity			

The results of the SEaTH analysis are presented in Table 3-3. The JM values for both datasets were close to 2, indicating a good separability between forest and canopy gaps. The respective separability threshold values for the CHM (9.965) and intensity (9.070) were subsequently used to build decision rules for classifying canopy gaps. A mean CHM value less than the threshold value (9.965) indicates canopy gap, whereas a mean CHM value greater than the threshold value indicates forest. Conversely, a mean intensity value greater than the threshold value (9.070) indicates canopy gap, whereas a mean intensity value less than the threshold value indicates forest. Consequently, for the combined dataset (i.e. CHM and intensity raster) a mean CHM value less than the threshold value (9.965) and mean intensity value greater than the threshold value (9.070) indicates canopy gap.

Table 3-3 Jeffries-Matusita distance (J-M) and separability thresholds for differentiating forest and canopy gaps.

Features	JM	Threshold	Notes
Mean CHM	1.996	9.965	The CHM best discriminates canopy gaps from forest where Mean CHM is less than 9.965.
Mean Intensity	1.810	9.070	The intensity raster best discriminates canopy gaps from forest where Mean intensity is more than 9.070.

A subset of the classification results for compartment F1 is shown in Figure 3-2. When comparing the classification output between Figure 3-2a and Figure 3-2b, for Figure 3-2b one can see that the waterbody (as indicated by the red arrow) has been excluded from the class canopy gaps and included in the general forest class. Similarly, for the combined dataset (Figure 3-2c), the waterbody has been excluded from the class canopy gaps. As explained in Table 3-3, the classified canopy gaps using the CHM (Figure 3-2a) are image objects where mean CHM is less than 9.965. Objects with mean CHM above 9.965 were classified as forest. Canopy gaps classified using the intensity raster (Figure 3-2b), were objects where mean intensity is more than 9.070. Subsequently, all objects having mean intensity less than 9.070 were classified as forest. For the combined dataset (Figure 3-2c), objects are classified as canopy gaps where both threshold criteria are met as described in Table 3-3. So, for an object to be classified as a canopy gap, mean CHM < 9.96511 AND mean intensity > 9.06973.

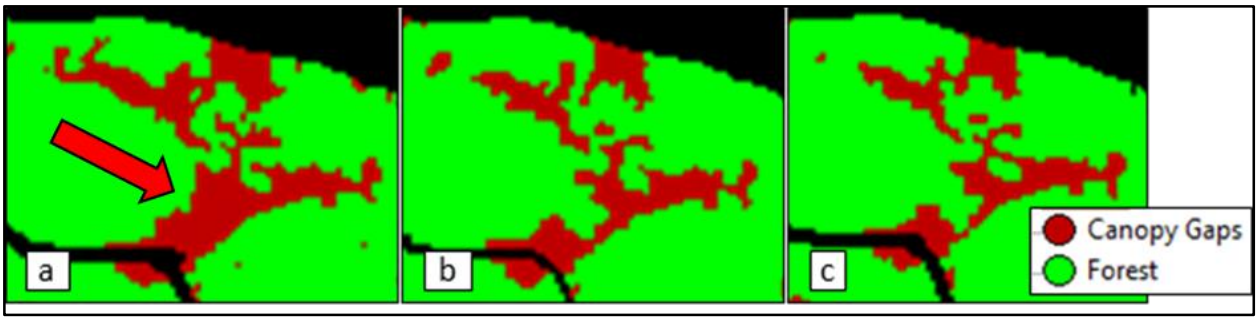


Figure 3-2 Subset of CHM (a), intensity (b), and combined (c) classified datasets of compartment F1.

Table 3-4 provides a summary of the rule-based classification accuracies for compartment F1 (used for model building) and compartment F3a (used for model testing). For compartment F1, overall accuracies ranged from 96% to 99%, with KHAT values ranging from 0.88 to 0.96. The results show that for compartment F1, using the combined dataset (i.e. CHM and intensity raster) yielded the highest overall accuracy (99%) and KHAT value (0.96), whereas using the intensity raster alone yielded the lowest overall accuracy (96%) and KHAT value (0.88). For compartment F3a, again, using the combined dataset yielded the highest overall accuracy (98%) and KHAT value (0.91). However, using the CHM alone yielded the lowest overall accuracy (94%) and KHAT value (0.82). For compartment F1, the lower accuracy obtained by the intensity raster is attributed to instances where the SEaTH intensity threshold (9.070) was exceeded. Similarly, lower accuracies were obtained using the CHM for compartment F3a, where the SEaTH height threshold (9.965) was exceeded. The results indicate that overall the highest thematic accuracy for canopy gap classification was obtained using a combination of the CHM and intensity raster.

Table 3-4 Thematic accuracy assessment for compartment F1 and compartment F3a.

	F1			**F3a**		
	CHM	Intensity	Combined	CHM	Intensity	Combined
Overall accuracy (%)	98	96	99	94	96	98
KHAT	0.92	0.88	0.96	0.82	0.87	0.91

The results of the area-based assessment are shown in Figure 3-3. For compartment F1, accuracies ranged from 75% to 95%, whereas for compartment F3a, accuracies ranged from 79% to 92%. Similar to the thematic accuracy results in Table 3-4, using the combined dataset (CHM and intensity raster) yielded the highest areal accuracies for compartment F1 (used for model building) and compartment F3a (used for model testing). When assessing the accuracy of canopy gaps using area as opposed to points, using both CHM and intensity information is particularly useful.

Additionally, a visual interpretation of the data showed that canopy gaps extracted using the combined dataset overlapped better with the reference data (i.e. aerial imagery), compared with canopy gaps extracted using the CHM and intensity raster alone.

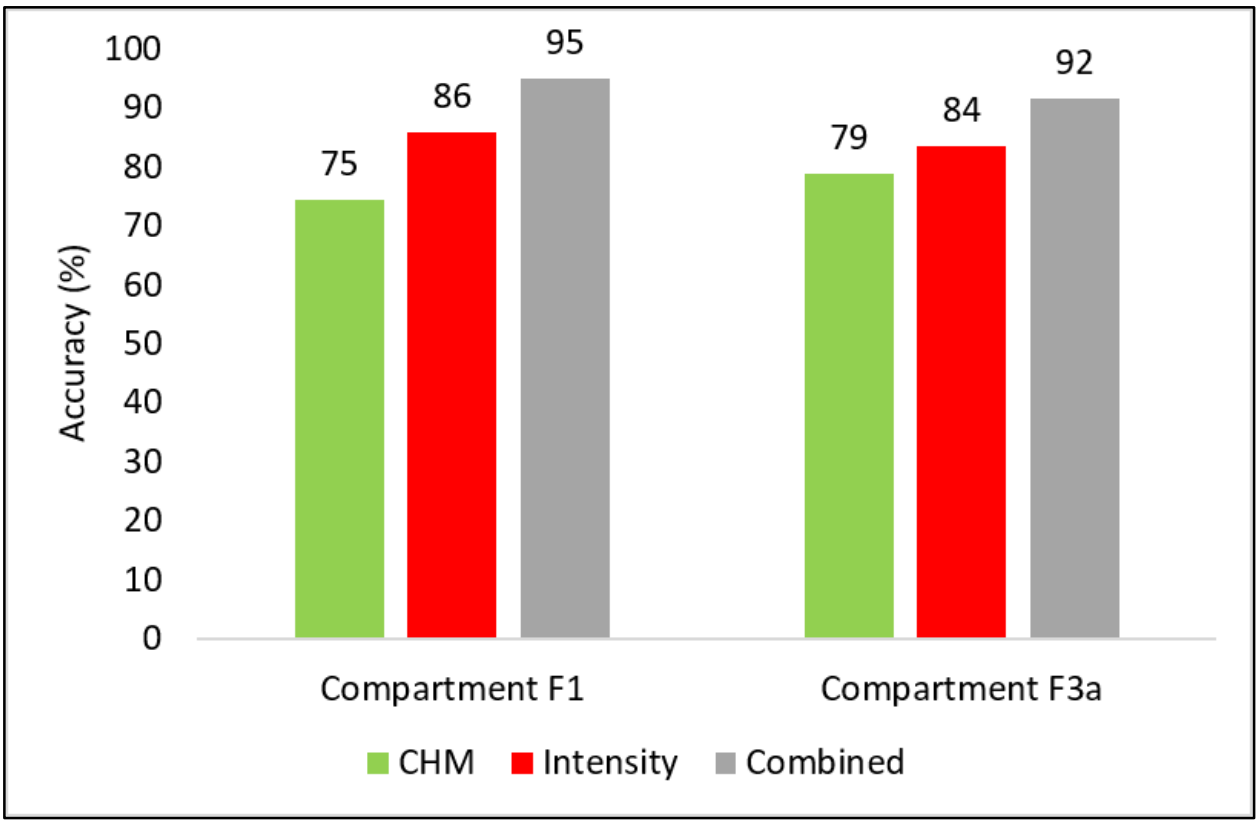


Figure 3-3 Comparative area based assessment for compartment F1 and compartment F3a.

3.4.2 Assessing spatial clustering using Getis-Ord Gi*

The results for the hotspot analysis for block E, block F, and the combined blocks, is presented in Figure 3-4. The distance band used for block E, block F, and the combined block was 300, 400, and 1700 respectively. In block E (Figure 3-4a), approximately 22% (n = 13) of canopy gaps (n = 58) were identified as hotspots (spatial clustering), with z-scores ranging from 1.65 to 2.58 and p-values ranging from 0.009918 to 0.099580. In contrast to block E, block F (Figure 3-4b) only showed spatial clustering of two canopy gaps with both having a z-score of 4.50 and p-value of 0.000007.

An assessment of spatial clustering in the combined block showed that of the total number of canopy gaps (n = 196), only eight canopy gaps displayed spatial clustering (hotspot). All of these hotspots (n = 8) had the same z-score (1.75) and p-value (0.080694). However, spatial clustering in the combined block appeared fragmented compared with spatial clustering in block E and block F. We attribute this to the increased scale of analysis in the combined block.

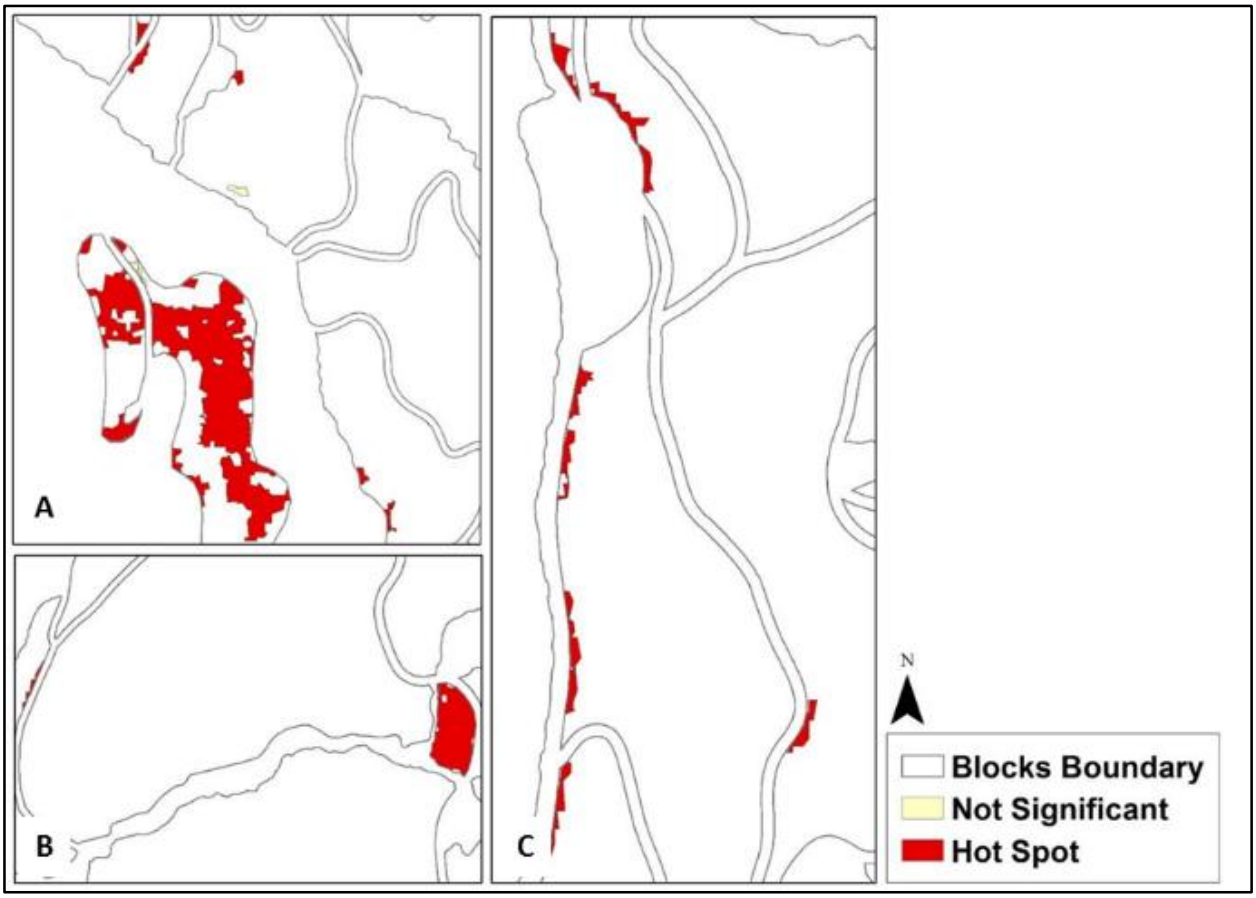


Figure 3-4 Identified hotspots in block E (A), block F (B), and the combined block (C).

3.4.3 Spatial characterisation of canopy gaps using FRAGSTATS

Table 3-5 shows the results of the canopy gap analysis using FRAGSTATS. The higher PD value (125.39) and PLAND value (3.02) for block F indicated a higher number of canopy gaps comprising a larger percentage of the landscape, compared with the PD value (42.23) and PLAND value (1.27) for block E. Additionally, the LSI was significantly higher for block F (20.57) compared with block E (11.84). The higher LSI indicated a higher dispersion of canopy gaps, coupled with a higher shape irregularity. The larger LSI for the combined block (23.55) is attributed to the increased dispersion and irregularity of canopy gaps. The higher shape index for block F (1.84) indicated the presence of larger patches with higher irregularity, compared with block E (1.79). The shape index for the combined block (1.82) represents the average size and regularity of canopy gaps in block E and block F.

Table 3-5 FRAGSTATS metrics for canopy gaps at class level for block E, block F and the combined block (E + F).

Metrics	E Block	F Block	Combined block
Class level			
PLAND	1.27	3.02	2.02
PD	42.23	125.39	77.93
LSI	11.84	20.57	23.55
Shape Index	1.79	1.84	1.82

3.5 DISCUSSION

This study evaluated the use of LiDAR derived variables, specifically a CHM and intensity raster, to identify and delineate canopy gaps within a commercial *Eucalyptus grandis* forest. The analysis was undertaken within an object-based image analysis environment. Additionally, a statistical analysis assessed the spatial characteristics of the canopy gaps using Getis-Ord Gi* and FRAGSTATS. The following sections discuss the results in more detail.

3.5.1 Identifying and delineating canopy gaps

The accurate delineation of canopy gaps is important to forest managers, as it forms the basis for further research and management applications such as gap dynamics and silvicultural applications (Muscolo et al. 2014; Bonnet et al. 2015). Additionally, forest gap dynamics has been recognized as an important process in stand development (Kucbel et al. 2010). Different tree species tolerate shaded conditions differently, therefore, leading to changes in composition and competition of trees in the understory. In particular, larger canopy gaps increase the occurrence of shade-intolerant species and thus may lead to changes in stand composition (de Römer et al. 2007).

Bonnet et al. (2015) note that without reliable methods for accurate canopy gap mapping, the spatial pattern of canopy gaps may be misrepresented. The authors further note that analysing the spatial characteristics of canopy gaps underpins the understanding of forest regeneration, ecosystem dynamics, species diversification, and species distribution. The use of remote sensing data and techniques presents an accurate and cost-effective approach to broad-scale mapping of canopy gaps.

In this study, we employed an object-oriented approach to delineate canopy gaps using a LiDAR-derived CHM and intensity raster. Canopy gap classification using the combined CHM and intensity raster yielded thematic accuracies of 99% (KHAT = 0.96) and 98% (KHAT = 0.91) for the training and testing datasets respectively, and areal accuracies of 95% and 92% for the training and testing datasets respectively. The results of this study show a significant improvement on the results achieved by Gaulton and Malthus (2010) who achieved an overall accuracy of 85% (KHAT

= 0.66) using a CHM only for canopy gap delineation. Similarly, the rule-based approach used in this study yielded significantly higher overall accuracies compared with Bonnet et al. (2015) who used a per-object based supervised classification, using a CHM, canopy porosity index, and intensity raster. The authors reported an overall accuracy of 80%. The results of this study compare favourably to the results achieved by Vepakomma et al. (2008), who also used a CHM in an object-based approach to delineate canopy gaps. These authors reported an overall accuracy of 96%. The results obtained in this study thus highlight the efficiency and robustness of using LiDAR-derived variables, in particular, using a combined CHM and intensity raster, for detecting and delineating canopy gaps using an OBIA approach.

3.5.2 Spatial clustering and characterization of canopy gaps

Hotspot analysis is a useful analysis tool that enables the tracking of local concentrations of a feature, such as canopy gaps (Reddy et al. 2016). The importance of delineating canopy gaps ranges from understanding forest regeneration from an ecological perspective to estimating the amount of volume loss within a managed plantation (Barbati et al. 2009). Reddy et al. (2016) used hotspot analysis to track canopy gap hotspots caused by forest deforestation. Similarly, Barbati et al. (2009) performed hotspot analysis (Getis-Ord Gi*) to detect canopy gaps.

We used the Getis-Ord Gi* statistic (hotspot analysis) to assess spatial clustering of canopy gaps on a block level (i.e. block E and block F) as well as on a combined block level (i.e. block E + block F). Getis-Ord Gi* statistic was undertaken on the combined block to investigate spatial clustering of canopy gaps on the entire extent of E. grandis within the plantation. In block E, canopy gap hotspots reported z-scores ranging between 1.65 and 2.58, while z-scores of 4.50 were reported for canopy gap hotspots in block F. The combined block reported z-scores of 1.75 for all canopy gap hotspots.

Spatial clustering of approximately 22% of canopy gap hotspots were found in block E, while only approximately 2% of spatial clustering of canopy gap hotspots were reported in block F. In block E, a large canopy gap had a higher likelihood of being surrounded by similar sized canopy gaps, whereas canopy gaps in block F had a lower likelihood of having similar sized neighbouring canopy gaps. This resulted in a lower spatial clustering of canopy gap hotspots for block F. The combined block also reported a low spatial clustering of canopy gap hotspots (approximately 4%). The larger scale of analysis and larger varied sizes of canopy gaps at the combined block level resulted in a lower percentage of canopy gap hotspots compared with spatial clustering at individual block level (i.e. block E).

The spatial configuration of canopy gaps is generally influenced by the nature of the disturbance (Andrew et al. 2016). Canopy gaps caused by natural disturbances, e.g. wind, have differences in shape compared with canopy gaps caused by artificial disturbances such as harvesting (Andrew et al. 2016). Canopy gaps caused by natural disturbances have been reported to have varied shapes, such as dumb-bell, or triangular (Schliemann and Bockheim 2011; Muscolo et al. 2014), whereas the shape of a harvested canopy gap tends to be elliptical, square, or circular (Schliemann and Bockheim 2011; Muscolo et al. 2014). Within a commercial forest environment, tree felling usually results in larger canopy gaps due to a large number of trees being felled within a small area (Muscolo et al. 2014). To maximize timber yield, larger trees are felled, which also contribute to the increased canopy gap area (Sapkota and Odén 2009).

Vehmas et al. (2011) used several FRAGSTATS metrics to model canopy gaps within semi natural and managed forests. In this study, we used four FRAGSTATS statistical metrics, namely PLAND, PD, LSI, and shape index to spatially characterise canopy gaps within an *E. grandis* commercial forest. When comparing the results of block E with block F, we found that the patch density (PD) of canopy gaps for block F (125.39) was significantly higher than PD for block E (42.23). This significant difference resulted in a larger PLAND for block F (3.02) compared with block E (1.27). The LSI and shape index for block F (LSI = 20.57; shape index = 1.84) was also higher compared with block E (LSI = 11.84; shape index = 1.79), which indicates a higher occurrence of irregularly shaped canopy gaps coupled with higher dispersion of irregular canopy gaps in block F.

The results for the combined block indicate that canopy gaps within *E. grandis* occupy less than 3% of the total landscape (PLAND = 2.02). A PD of 77.93 and shape index of 1.82 indicates that the shape of canopy gaps are generally irregular. Additionally, the landscape shape index (LSI = 23.55) indicates that canopy gaps are more irregularly shaped and have a high level of dispersion. The results of the combined block is consistent with the results obtained for the individual blocks, i.e. block E and block F.

The results indicate that larger or more frequent disturbances occurred across block F compared with block E. Canopy gaps within both blocks were generally found to be irregularly shaped, which suggest that natural disturbances (e.g. tree mortality) occurred, in addition to management activities. A similar shape index (1.82) for the combined block suggests that the nature of disturbances was also natural, in addition to management activities. The PLAND for the combined block (2.02) compared with the higher PLAND for block F (3.02) and lower PLAND for block E (1.27) is expected, as PLAND for the combined block represents the average percentage of canopy gap area across both blocks. Additionally, the larger LSI for the combined block (23.55) compared

with block E (11.84) and block F (20.57) is expected, given the larger dispersion and greater variability in canopy gap shape.

Ultimately, these metrics (PLAND, PD, LSI, and shape index) should assist foresters and forest managers in quantifying their landscape, and in particular, canopy gaps within commercial plantations (McGarigal et al. 2012). This information can be used to improve the understanding of the mechanisms that influence canopy gap formation, i.e. natural or harvesting, and be integrated into existing and future forest management frameworks.

3.5.3 An operational framework for canopy gap modelling

Passive remote sensing systems have been widely explored for forestry applications. However, the ability of LiDAR to directly measure the ground and forest elevation by penetrating small canopy gaps in the forest canopy has attracted attention in forestry (Zhang 2008; Vepakomma et al. 2011). LiDAR has subsequently been exploited for both natural and commercial forestry applications. Using 3-Dimensional point data, LiDAR generates canopy surface, terrain, and canopy height models that can be used to identify and delineate canopy gaps (Kent et al. 2015).

The usefulness of OBIA is particularly significant when features are analysed using high resolution datasets such as LiDAR (Blaschke 2010). For example, when the size of the features of interest are small (i.e. a canopy gap), OBIA can be used to generate image objects to encompass the features of interest (Einzmann et al. 2017).

This study has successfully demonstrated the utility of LiDAR-derived variables for modelling canopy gaps within an OBIA environment. However, the sample size (n = 26) in respect of canopy gaps used to assess classification (thematic and area-based) accuracy is a limitation in this study. Future studies could consider analysing more than one plantation, or multiple species, although this may be infeasible. Additionally, a significant limitation using the intensity raster only is that the efficiency of discriminating between canopy gaps and forest diminishes in cases where low vegetation occurs within canopy gaps. Vegetation within canopy gaps results in canopy gaps having similar intensity values as forested areas. However, the results of this study show that this limitation is readily overcome using a combination of CHM and intensity. Thus, using both CHM and intensity information within a combined dataset, forest managers can potentially implement this methodology to detect canopy gaps and incorporate gap-based silviculture to improve stand composition and forest structure.

Further research is required to investigate the spatial characteristics of canopy gaps across *Eucalyptus* species as well as other commercial forest species. Additionally, the use of image

object attributes (Blaschke 2010; Bonnet et al. 2015; Einzmann et al. 2017) may assist in the classification of canopy gaps in more complex forest environments such as natural forests. Forest managers would, therefore, be provided with information to better understand canopy gap dynamics across the entire plantation.

3.6 CONCLUSION

The primary aim of this study was to model canopy gaps within a commercial *Eucalyptus grandis* plantation using LiDAR derived variables, i.e. a CHM and an intensity raster. To the best of our knowledge, as consulted with the literature, this is probably the first study to attempt canopy gap delineation using LiDAR in South Africa. The results of this study highlight that using a combined CHM and intensity raster provides improved accuracies for modelling canopy gaps within a commercial *E. grandis* forest. The overall results show promise as a viable methodology that can be operationalised for modelling canopy gaps within a commercial forestry environment.

CHAPTER 4: INVESTIGATING THE UTILITY OF CANOPY GAPS AND FOREST CANOPY FOR DISCRIMINATING EUCALYPTUS SPECIES USING RF AND LIDAR

4.1 INTRODUCTION

The importance of forest species mapping and discrimination has been reported by a number of studies (Gjertsen 2007; Dalponte et al. 2008; Kim et al. 2009b; Pekkarinen et al. 2009; Corona 2010; Wolter and Townsend 2011; Dalponte et al. 2012; Peerbhay et al. 2013; Peerbhay 2014; Waser et al. 2015; Qin et al. 2016; Mulyani and Jepson 2017). The ability to discriminate forest species has both economic and conservation benefits (Kim et al. 2009b; Shang and Chrisholm 2014). Economically, having forest species information assists in estimating biomass and wood production and is important to develop growth and yield models (Ko et al. 2013). Additionally, species information enables estimating timber volume. Having this information is invaluable for commercial plantations (Dalponte et al. 2008). In South Africa, hardwoods are utilized for pulp and timber production. Timber production yields approximately R1.8 billion, whereas woodpulp and paper products produce approximately R11 billion (DAFF 2012).

From a conservation perspective, forest species mapping is important for the management of forest communities as well as promoting effective assessment of species vulnerability to threats (such as pests or droughts) (Hill et al. 2010; Shang and Chrisholm 2014; Abdollahnejad et al. 2017). Furthermore, forest species mapping enables biodiversity maintenance, stem volume estimation, sustainable forest management, forest disturbance detection, as well as habitat mapping (Barilotti et al. 2009; Falkowski et al. 2009; Immitzer et al. 2012; Gosh et al. 2014; Waser et al. 2015).

Traditionally, field surveys and aerial photograph interpretation were some of the main approaches to acquire information about forest species (Aldred and Hall 1975; Immitzer et al. 2012). These methods, however, are costly, labour intensive and time consuming (Van Coillie et al. 2007; Kim et al. 2009a; Cho et al. 2012; Peerbhay et al. 2013). Remote sensing has been utilized since 1970 to aid forestry practices. It enables the classification of vegetation types whilst being more cost-effective and less labour-intensive than traditional in situ methods and aerial photograph interpretation (Bradley and Fleishman 2008; Cho et al. 2012). Remote sensing is able to map large forested areas at higher spatial resolutions (Xie et al. 2008; Immitzer et al. 2012). In addition to high spatial resolution, passive remote sensing sensors have a number of spectral bands to assist in more precise tree species discrimination (Arenas-Castro et al. 2013). Subsequently, a number of studies have utilized passive sensors to classify forest species (Everitt et al. 2008; Mallinis et al. 2008; Kim et al. 2009a; Hill et al. 2010; Arenas-Castro et al. 2013; Li et al. 2013b; Wang et al.

2016).

Mallinis et al. (2008) used Quickbird imagery to classify dominant forest vegetation in North Greece using an object-based approach. The authors utilized a Fractal Net Evolution image segmentation prior to employing three classification methods. The classifiers include: nearest neighbour, classification trees, as well as a combination of classification trees and local indicators of spatial association (texture features). The highest overall accuracy and KHAT were obtained using classification trees with additional texture features (i.e. overall accuracy = 78.11% and KHAT = 0.75). Similar to Mallinis et al. (2008), Arenas-Castro et al. (2013) used Quickbird imagery for forest vegetation mapping in Sierra Morena, southern Spain.

Arenas-Castro et al. (2013) discriminated wild pear trees (*Pyrus bourgaeana*) from other woodland vegetation using maximum likelihood classification with an overall accuracy of 80.42% and KHAT of 0.78 for species discrimination. More recently, Abdollahnejad et al. (2017) also used Quickbird imagery to discriminate dominant tree species in Gorgan city, Iran. The authors tested the utility of three ensemble learning classifiers (i.e. random forest (RF), support vector machines (SVM) and k-nearest neighbour (*k*-NN). Of the three classifiers, RF was the most efficient in discriminating *Fagus orientalis*, *Carpinus betulus*, *Parrotia persica* and various species with an overall accuracy of 63.85%.

Very high spatial resolution passive sensors can discriminate forest species more efficiently compared with low and medium spatial resolution sensors; however, similarity of species spectral reflectance is still problematic (Lucas et al. 2008; Hill et al. 2010; Korpela et al. 2010). Active sensors such as light detection and ranging (LiDAR) overcomes these obstacles and are capable of discriminating forest species more efficiently compared with passive sensors (Ke et al. 2010). This is achieved by providing three dimensional and species-specific structural information (Ke et al. 2010; Kim et al. 2011a). Subsequently, several studies have utilized LiDAR for species mapping (Brandtberg 2007; Kim et al. 2009b; Ørka et al. 2009; Vauhkonen et al. 2009; Korpela et al. 2010; Vaughn et al. 2012; Li et al. 2013a; Yu et al. 2014; Cao et al. 2016).

In Seattle, North America, Kim et al. (2009b) classified seven coniferous and eight broadleaved tree species using two LiDAR datasets (i.e. one for leaf-on and one for leaf-off conditions). Using intensity features and a linear discriminant function, the authors reported accuracies of 83.4% and 73.1% for leaf-off and leaf-on datasets, respectively. The authors tested using a combination of both datasets and obtained an improved overall classification accuracy of 90.6%. The utility of LiDAR intensity for species discrimination was also investigated by Ørka et al. (2009). The authors classified coniferous and deciduous tree species in south-eastern Norway. Additionally,

the authors generated structural features (i.e. normalized height features, canopy penetration depth, other height features, and crown density features) to assist in species classification using a linear discriminant analysis. Overall classification results using intensity features yielded an accuracy of 70%, whereas classification using structural information (i.e. crown density features) obtained an accuracy of 77%. The authors also used a combination of features and obtained improved results with an overall accuracy of 88%.

Vauhkonen et al. (2009) and Li et al. (2013a) tested the utility of texture information for forest species discrimination. In addition to texture features, Vauhkonen et al. (2009) derived tree crown approximations such as alpha shape, height, and intensity to discriminate between Scandinavian commercial species (i.e. pine, spruce, and deciduous species) in southern Finland. Using a discriminant analysis classifier and a combination of intensity and texture features, the authors obtained an overall classification accuracy of 91% and KHAT of 0.84. Assessing the utility of texture features and tree crown characteristics to classify four forest species (i.e. sugar maple, trembling aspen, jack pine, and eastern white pine) in Ontario, Canada was undertaken by Li et al. (2013a). More specifically, the LiDAR-derived features include a three-dimensional texture, relative degree of foliage clustering, relative scale of foliage clustering, and gap distribution within tree crowns. Using a linear discriminant analysis, the authors reported a classification accuracy of 77.5% and KHAT of 0.7.

Several studies have successfully used LiDAR and the RF classifier for species discrimination (Korpela et al. 2010; Yu et al. 2014; Cao et al. 2016). For example, Korpela et al. (2010) undertook tree species classification to discriminate Scots pine, Norway spruce, and birch in southern Finland. The authors derived various intensity features and utilized three classifiers, including linear discriminant analysis, *k*-NN, and RF. Using a combination of two discrete LiDAR datasets, overall classification accuracies ranged from 89.4% to 90.8%, with RF obtaining an improved overall accuracy and KHAT of 90.8% and 0.84, respectively. Similarly, Yu et al. (2014) also utilized RF to discriminate between Scots pine, Norway spruce and birch in Evo, southern Finland. The authors assessed the efficiency of using a combination of full-waveform and discrete LiDAR features. These features include mean heights, standard deviation of heights, and a mean of full-waveform data interacting with a tree, among others. The authors reported that using a combination of both waveform and discrete LiDAR data features yielded the highest classification accuracy, with an overall accuracy of 73.4%. In Changshu, southeast China, Cao et al. (2016) utilized RF to discriminate six forest species using full-waveform LiDAR. Forest species include Masson pine, Chinese fir, Slash pines, Sawtooth oak, Sweet gum, and Chinese holly. Overall accuracy and KHAT for species classification were 68.6% and 0.62, respectively. An improved overall accuracy

and KHAT were obtained when undertaking species classification on four main species (overall accuracy = 75.8% and KHAT = 0.68).

In South Africa, *Eucalyptus* tree species are important for commercial forestry. Of the *Eucalyptus* species, *Eucalyptus grandis* is the dominant commercial hardwood specie, accounting for approximately 48% of all total hardwood area (DAFF 2012). Discriminating *Eucalyptus* species effectively using remote sensing would, therefore, be beneficial to commercial forestry. Sensors, such as LiDAR have been shown to possess capability to obtain sufficient species discrimination results (Lucas et al. 2008; Ke et al. 2010; Shang and Chrisholm 2014). LiDAR intensity and texture features have also been reported to be beneficial in forest species discrimination (Vauhkonen et al. 2009). Furthermore, the efficiency of the RF classifier for forest species discrimination using LiDAR data has also been shown (Korpela et al. 2010; Yu et al. 2014).

A large body of literature focused specifically on extracting forest canopy information for species classification with LiDAR data and RF. To date, only one study looked at forest gaps (Li et al. 2013a). However, Li et al. (2013a) looked at gaps between tree crowns. The authors evaluated individual tree structure and tree crown information derived from LiDAR to classify species. More specifically, Li et al. (2013a) investigated several LiDAR features; including three-dimensional texture, foliage clustering degree, foliage clustering scale, and gap distribution of individual trees. Based on the gap variability of intra tree crown gaps as investigated by Li et al. (2013a), this study investigates the variability of intra tree species canopy gaps. More specifically, this study investigates the use of canopy gaps for discriminating *Eucalyptus grandis* and *Eucalyptus dunnii* in a commercial plantation, South Africa using LiDAR-derived intensity and texture features and the RF classifier.

4.2 MATERIALS AND METHODS

4.2.1 Study Area

The study area is the Sappi Riverdale plantation, comprising a total plantation area of 5999ha (Figure 4-1). The plantation is located near Richmond in KwaZulu-Natal, South Africa. The *Eucalyptus* plantation comprises three species, namely, *E. grandis*, *E. smithii*, and *E. dunnii*. The species of interest are *E. grandis* and *E. dunnii*, consisting of n = 15 and n = 10 compartments, respectively (Macfarlane 2006). The species and ages of each compartment can be seen in Table 4-1.

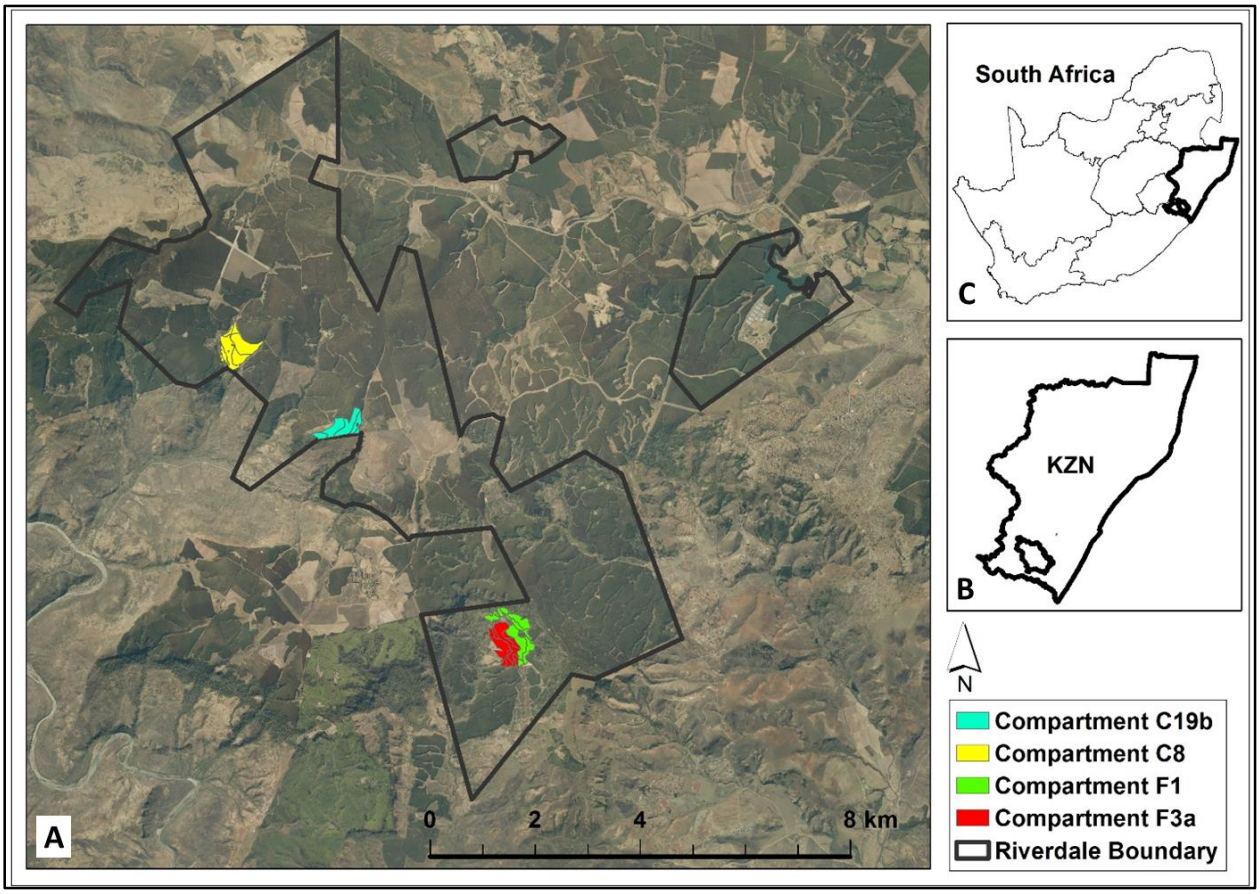


Figure 4-1 The Sappi Riverdale plantation is (a) located in KwaZulu-Natal (b), South Africa (c). Background image is ESRI ArcGIS online's 50 cm colour imagery for South Africa.

Table 4-1 Species description and age per compartment.

Compartment	Age (years)	Species
C19b	2.48	*E. dunnii*
C8	5.24	*E. dunnii*
F1	4.81	*E. grandis*
F3a	4.38	*E. grandis*

4.2.2 LiDAR and field data

All LiDAR and field data were supplied by Sappi Forests. LiDAR data were surveyed using a Leica ALS50-2 scanner (Table 4-2). Field data was supplied in the form of enumerated plot data for each compartment. Four compartments were selected for the analysis of this study. Two of these compartments contain *E. grandis* (compartments F1 and F3a), whereas the remaining two comprise *E. dunnii* (compartments C19b and C8). In addition to using canopy gaps alone for species classification, further investigation was undertaken by using a combination of forest canopy and canopy gap LiDAR-derived features. This was comparatively to using only canopy gap features, for species discrimination. For each compartment, intensity and texture information from both forest canopy and canopy gaps were extracted for the analysis. However, limited canopy gaps occurred across all *Eucalyptus* compartments. The four selected compartments were

purposively sampled as they contained the highest number of canopy gaps. The total number of canopy gaps per compartment were, 68 canopy gaps within compartment F3a, 54 within compartment F1, 59 within compartment C19b, and 39 canopy gaps within compartment C8. To keep the number of canopy gaps consistent across all four compartments, 30 canopy gaps polygons were randomly selected per compartment. Additionally, 30 forest canopy polygons were also randomly selected per compartment. Forest canopy and canopy gap polygons per compartment were confirmed using the enumerated plot data, aerial imagery, and based on a field visit to the plantation. Aerial imagery of 15cm spatial resolution was acquired on 12 April 2014 at an average flying height of 213m. The aerial imagery was not used for any direct analysis due to an inherent spatial inaccuracy between the aerial imagery and LiDAR data.

Table 4-2 LiDAR data capture information conducted for the Sappi Riverdale plantation.

Number of returns per emitted LiDAR signal	4
Pulse rate (Hz)	1260 000
Scan rate (Hz)	53
Average flying height (m)	820
Survey period	15 to 22 March 2014

4.2.3 Intensity features

To assist in discriminating *E. grandis* and *E. dunnii*, various intensity features (measuring the return signal strength of transmitted LiDAR pulses) were calculated using FUSION/LDV v3.60 (Yunfei et al. 2008; Maltamo et al. 2014; FUSION 2016). Using command line programs, 34 descriptive statistics (features) of LiDAR intensity were calculated using a tool called GridMetrics (Table 4-3) (FUSION 2016). The output of the function is a CSV (comma separated values) comma delimited file containing all 34 intensity features for each LiDAR input cell. These features were calculated for each compartment. Subsequently, the CSV2Grid tool was utilized to convert each intensity feature in CSV format into ASCII grid format (FUSION 2016).

FUSION/LDV has previously been used by a number of studies (Kim et al. 2009b; Kim et al. 2011a; d'Oliveira et al. 2012; Bright et al. 2013). For example, Kim et al. (2009b) used FUSION/LDV to undertake tree species differentiation using LiDAR intensity data, whereas Kim et al. (2011a) used FUSION/LDV to assist in individual tree genera classification. d'Oliveira et al. (2012) used FUSION to derive a variety of LiDAR elevation metrics for estimating forest biomass and to identify low-intensity logging areas. FUSION/LDV was also utilized by Bright et al. (2013) to derive various LiDAR metrics to assist in predicting live and dead tree basal areas.

In this study, the 30 randomly selected forest canopy polygons and 30 randomly selected canopy

gap polygons were utilized to extract all intensity features for each compartment in ArcMap v10.3.1 (ESRI 2015).

Table 4-3 FUSION derived intensity features computed for each cell.

Feature	**Description**
Total return count above *htmin*	The total number of returned LiDAR signal above the minimum intensity for each cell
Minimum	The minimum intensity value for each cell
Maximum	The maximum intensity value for each cell
Mean	The mean intensity value for each cell
Mode	The mode intensity value for each cell
Standard deviation	The standard deviation intensity value for each cell
Variance	The variance intensity value for each cell
Coefficient of variance	The coefficient of variance intensity value for each cell
Interquartile distance	The interquartile distance intensity value for each cell
Skewness	The skewness intensity value for each cell
Kurtosis	The kurtosis intensity value for each cell
Average absolute deviation	The average absolute deviation intensity value for each cell
L-moments (L1 – L4)	The L-moments 1-4 intensity value for each cell
L-moment coefficient of variance	The L-moment coefficient of variance intensity value for each cell
L-moment skewness	The L-moment skewness intensity value for each cell
L-moment kurtosis	The L-moment kurtosis intensity value for each cell
P01 – P99	Percentile values 1 to 99 intensity value for each cell

4.2.4 Texture features

In addition to intensity features, 12 texture features were calculated for all four compartments in eCognition developer 9 (Trimble 2016) (Table 4-4). Prior to texture feature generation, a multiresolution segmentation (MRS) was undertaken to derive object features using a combined LiDAR canopy height model (CHM) and intensity raster. MRS is a region merging algorithm that derives image objects from pixels (Belgiu and Drăguţ 2014). Image objects are iteratively merged and determined by some homogeneity criteria (Rahman and Saha 2008). The homogeneity criteria consist of parameters that must be defined, such as scale, compactness, and shape (Drăguţ et al. 2010). Scale determines the size of resulting objects, whereas shape and compactness determines the overall shape and compactness of resulting objects (Definiens 2007; Rahman and Saha 2008).

Several grey-level co-occurrence matrices (GLCM) and grey-level difference vector (GLDV) texture features were calculated on the resulting MRS object features. Texture measures the differences in levels or grey tone of objects (Haralick et al. 1973). GLCM measures the spatial relationships of co-occurrence grey levels at specific distances and directions, whereas GLDV measures GLCM diagonals (Mhangara and Odindi 2013; Dian et al. 2015). For MRS, various scale, shape, and compactness values were tested. However, the best combination of MRS

parameters was in line with Lombard et al. (2017). The authors used a scale factor of 5, as well as shape and compactness values of 0.1 and 0.5, respectively. Similar to intensity feature extraction, 30 randomly selected canopy gap polygons and 30 randomly selected forest canopy polygons were utilized for texture feature extraction for each compartment in ArcMap.

Table 4-4 GLCM (all directions) and GLDV (all directions) texture features computed for each object.

Feature	**Description**
GLCM Angular 2nd moment	Grey-level co-occurrence matrix angular second moment value per object
GLCM Contrast	Grey-level co-occurrence matrix contrast value per object
GLCM Correlation	Grey-level co-occurrence matrix correlation value per object
GLCM Dissimilarity	Grey-level co-occurrence matrix dissimilarity value per object
GLCM Entropy	Grey-level co-occurrence matrix entropy value per object
GLCM Homogeneity	Grey-level co-occurrence matrix homogeneity value per object
GLCM Mean	Grey-level co-occurrence matrix dissimilarity value per object
GLCM Standard deviation	Grey-level co-occurrence matrix standard deviation value per object
GLDV Angular 2nd moment	Grey-level difference vector angular second moment value per object
GLDV Contrast	Grey-level difference vector contrast value per object
GLDV Entropy	Grey-level difference vector entropy value per object
GLDV Mean	Grey-level difference vector mean value per object

4.2.5 Species classification using Random Forest

Using the extracted intensity and texture features, a random forest (RF) classification was employed to discriminate *E. grandis* and *E. dunnii*. This was undertaken using the randomForest package in R version 3.4.1 (Liaw and Wiener 2002; R Development Core Team 2017). RF is an ensemble classifier that builds a large number of decision trees (*ntree*) (Breiman 2001). At each node split, a bootstrap sample (*mtry*) of the original data is selected (Breiman 2001; Miao et al. 2012; Chen et al. 2014). Subsequently, a majority vote is applied where *ntree* predictions are aggregated (Liaw and Wiener 2002).

In this study, default values were used for *ntree* and *mtry*, where *ntree* = 500 and *mtry* = $\sqrt{p}$. *Mtry* is the square root of the number of features used for node splitting within each tree (Belgiu and Drăguţ 2016). The value of *mtry* differed based on the number of input features used for RF classification. For example, *mtry* = $\sqrt{p(intensity)}$ = 5, *mtry* = $\sqrt{p(texture)}$ = 3, and *mtry* = $\sqrt{p(intensity + texture)}$ = 6. Additionally, 70% of the data were used for training, and the remaining 30% were used as the test set (Breiman 2001). The utility of RF for species classification were assessed using LiDAR intensity and texture features extracted from canopy gaps. Additionally, species discrimination was undertaken using a combination of forest canopy and canopy gaps. The RF model was undertaken using the following compartment combinations:

compartment C19b compared with compartment F3a, compartment C8 with compartment F1, compartment C19b with compartment F1, and compartment C8 with compartment F3a.

4.2.6 Accuracy Assessment

The RF model accuracy was assessed using training and test accuracies. For train accuracy, the out of bag (OOB) error estimate was utilized to estimate train classification accuracy using the train set in RF (Cao et al. 2016). Test accuracy was assessed using overall accuracy (OA). Additionally, the KHAT statistic was also utilized for both train and test accuracies. KHAT (Equation 4-1) assesses chance agreement against actual classification agreement and is defined according to the following formula (Congalton and Green 2009):

$$\widehat{K} = \frac{P_o - P_C}{1 - P_C} \qquad \text{Equation 4-1}$$

Where: $P_o = \sum_{i=1}^{k} P_{ii}$ is the actual agreement and $P_c = \sum_{i=1}^{k} P_{i+}\, P_{+j}$ is the chance agreement.

4.3 RESULTS

In this study, RF classifier was used to discriminate *E. grandis* and *E. dunnii* using LiDAR-derived intensity and texture features. The results can be seen in Table 4-5. The efficiency of species classification using varying compartment combinations was assessed i.e. compartment C19b and F3a, compartment C8 and F1, compartment C19b and F1, and compartment C8 and F3a. Using this approach, the influence of varying species ages on species discrimination could be assessed. Additionally, for each compartment, species classification was undertaken using only canopy gaps as well as a combination of forest canopy and canopy gaps.

Table 4-5 Species classification results between *E. grandis* and *E. dunnii* using canopy gap alone (G) as well as using a combination of canopy gaps forest canopy (G and F).

		C19b and F3a				**C8 and F1**				**C19b and F1**				**C8 and F3a**			
		Train		**Test**		**Train**		**Test**		**Train**		**Test**		**Train**		**Test**	
Extracted from	**Features**	**OOB error**	**KHAT**	**OA**	**KHAT**	**OOB error**	**KHAT**	**OA**	**KHAT**	**OOB error**	**KHAT**	**OA**	**KHAT**	**OOB error**	**KHAT**	**OA**	**KHAT**
G	Intensity & Texture (*n* = 46)	**7.89**	**0.84**	**90.91**	**0.81**	30.23	0.39	64.71	0.29	**8.89**	**0.82**	**86.67**	**0.73**	19.05	0.61	77.78	0.5
G	Intensity (*n* = 34)	21.43	0.57	77.78	0.54	**28.57**	**0.43**	**66.67**	**0.33**	29.17	0.42	66.67	0.25	**17.95**	**0.64**	**80.95**	**0.61**
G	Texture (*n* = 12)	13.04	0.74	85.71	0.71	38.1	0.24	61.11	0.22	12.5	0.75	85	0.7	34.21	0.28	59.09	0.14
F and G	Intensity & Texture (*n* = 46)	**3.66**	**0.92**	**94.74**	**0.88**	28.4	0.43	71.79	0.43	**8.14**	**0.84**	**91.18**	**0.82**	16	0.68	77.78	0.56
F and G	Intensity (*n* = 34)	19.1	0.62	80.65	0.61	**26.19**	**0.48**	**72.22**	**0.45**	29.11	0.42	68.29	0.36	**13.64**	**0.73**	**84.38**	**0.69**
F and G	Texture (*n* = 12)	15.48	0.69	83.33	0.67	30.68	0.39	65.62	0.31	8.7	0.83	85.71	0.71	27.27	0.46	68.75	0.37

4.3.1 Species classification using canopy gaps and a combination of forest canopy and canopy gaps

The results of species classification using canopy gaps and a combination of forest canopy and canopy gaps can be seen in Table 4-5. Species classification using canopy gap features for compartment C19b and F3a yielded OOB errors ranging from 21.43 to 7.89. Test accuracies ranged from 77.78% to 90.91%, with KHAT values from 0.54 to 0.81. The results for species discrimination using compartments C19b and F3a show that, using a combination of intensity and texture features yielded the lowest OOB error (7.89) and highest test accuracy (OA = 90.91% and KHAT = 0.81). However, using intensity features exclusively yielded the highest OOB error (21.43) and lowest test accuracy (OA = 77.78% and KHAT = 0.54). Similarly, for compartment C19b and compartment F1, using both intensity and texture features yielded improved species discrimination results, with a OOB error of 8.89 and test accuracy of 86.67% (KHAT = 0.73).

Species classification using compartments C8 and F1 yielded OOB errors ranging from 38.1 to 28.57. Test accuracies and KHAT values ranged from 61.11% to 66.67% and 0.22 to 0.33, respectively. For these compartments, using intensity features alone yielded improved results, with a OOB error of 28.57 and test accuracy of 66.67% (KHAT = 0.33). In contrast, using texture yielded the highest OOB error (38.1) lowest test accuracies (OA = 61.11 and KHAT = 0.22). Similar to compartments C8 and F1, compartments C8 and F3a also yielded improved results using intensity features, with a OOB error of 17.95. Additionally, classification test accuracy and KHAT values using intensity were 80.95% and 0.61, respectively.

An evaluation of the results for species classification using forest canopy and canopy gaps for compartments C19b and F3a obtained OOB errors ranging from 19.1 to 3.66. Test accuracies ranged from 80.65% to 94.74%, with KHAT values ranging from 0.61 to 0.88. For these compartments, using a combination of intensity and texture features for species classification obtained the lowest OOB error (3.66) and highest test accuracies (OA = 94.74% and KHAT = 0.88). Similarly, for compartments C19b and F1, using a combination of intensity and texture features yielded improved results. The OOB error of the combined features was 8.14. Test accuracy was 91.18%, and KHAT was 0.82 using intensity and texture features.

Compartments C8 and F1 yielded OOB errors varying from 30.68 to 26.19. Test accuracies ranged from 65.62% to 72.22%, with KHAT values ranging from 0.31 to 0.45. In contrast to compartment C19b and F3a, compartment C8 and F1 obtained improved results using intensity features. The OOB error using intensity was 26.19. Test accuracy and KHAT using intensity was 72.22% and

0.45, respectively. Compartments C8 and F3a also obtained improved species discrimination results using intensity information within both canopy gaps and forest canopy. A OOB error of 13.64 was obtained, whereas test accuracy and KHAT were 84.38% and 0.69, respectively.

The mean trees/ha and stumps/ha for each compartment were obtained using enumerated plot data and can be seen in Table 4-6. From the data in Table 4-6, it is apparent that compartments C8 and F1 have a larger difference in mean trees/ha (24) and mean stumps/ha (324) compared with compartments C19b and F3a (difference in mean trees/ha = 16 and mean stumps/ha = 110). This indicates that compartments C8 and F1 have a larger variation in mean trees/ha and mean stumps/ha. The results suggest that compartments with an increased variation in mean trees/ha and mean stumps/ha (i.e. compartments C8 and F1) yield improved classification results using intensity features.

Table 4-6 Mean tree height, mean trees/ha, and mean stumps/ha of each compartment.

Variables	**C19b**	**F1**	**C8**	**F3a**
Mean tree height (m)	10.08	21.92	15.95	21.05
Mean trees/ha	1546	1467	1491	1562
Mean stumps/ha	1444	1402	1078	1554

Similar to compartments C8 and F1, compartments C8 and F3a also yielded improved species classification results using intensity features. Although a smaller difference in mean trees/ha (71) for these compartments can be seen compared with compartments C19b and F1 (mean trees/ha = 79), this minor difference is offset by a significantly larger difference in mean stumps/ha (476) (mean stumps/ha for compartments C19b and F1 = 42).

4.3.2 The influence of varying ages for discriminating *E. grandis* and *E. dunnii*

The results of this study show that species classification using RF and LiDAR intensity features and texture features can be used to obtain efficient *Eucalyptus* species classification results. However, using a combination of intensity and texture features yields improved results, specifically for compartments with an increased age difference (i.e. compartments C19b and F3a and compartments C19b and F1). For example, using a combination of forest canopy and canopy gaps and intensity and texture features, compartments C19b and F3a obtained a test accuracy of 94.74% (age = 1.9 years), whereas compartments C8 and F1 obtained a test accuracy of 71.79% (age = 0.43 years). The results suggest compartments with an age difference of more than one year (i.e. compartments C19b and F1 and compartments C19b and F3a) yield more efficient results using a combination of intensity and texture features, whereas compartments with an age difference of less than one year (i.e. compartments C8 and F1 and compartments C8 and F3a) yield

improved results using intensity features alone.

Furthermore, compartments with an age difference of more than one year obtained improved results using texture features compared with intensity features. For these compartments, differences in tree characteristics such as tree height (Table 4-6) caused by increased age differences result in varying textures (Kayitakire et al. 2006). Therefore, yielding improved species discrimination using texture information. Similarly, using canopy gaps for compartments with an increased age difference obtained improved species classification results using texture features. Younger compartments (i.e. C19b) generally have bare canopy gaps, whereas older compartments (i.e. F1 and F3a) have sparse or dense vegetation (such as sprouts or shrubs) within canopy gaps. This results in varying textures and improved species discrimination using texture. Additionally, during field visits, it was confirmed that significant height differences of forest canopy and vegetation within canopy gaps occur between varying species ages.

4.4 DISCUSSION

This study evaluated the potential to discriminate *Eucalyptus* species within a commercial plantation using LiDAR intensity and texture features. Specifically, a random forest (RF) classifier was used to classify two *Eucalyptus* species, i.e. *E. grandis* and *E. dunnii*. Additionally, an assessment was undertaken to classify species using intensity and texture information contained within canopy gaps as well as information contained within both forest canopy and canopy gaps.

4.4.1 Species classification using canopy gaps

Forest species classification has traditionally been undertaken using ground-based methods and aerial imagery (Donoghue et al. 2007; Dalponte et al. 2008; Puttonen et al. 2010). However, more accurate and precise forest species information is invaluable for commercial forestry as well as conservation sectors (Moffiet et al. 2005; Puttonen et al. 2010). Recent literature has reported sufficient accuracies for forest species discrimination using LiDAR data, specifically using information from forest canopy (Ørka et al. 2009; Vauhkonen et al. 2009; Korpela et al. 2010; Vaughn et al. 2012; Yu et al. 2014; Cao et al. 2016). This study sought to accurately classify *E. grandis* and *E. dunnii* using LiDAR intensity and texture information contained within canopy gaps. Additionally, intensity and texture information from both forest canopy and canopy gaps was also evaluated for species discrimination.

Forest species discrimination using RF and a combination of LiDAR intensity, texture features and canopy gaps obtained a OOB error of 7.89 and test accuracy of 90.91% (KHAT = 0.81). Korpela et al. (2010), Yu et al. (2014) and Cao et al. (2016) also utilized a RF algorithm and

LiDAR data for species discrimination. However, all three studies used forest canopy information for classification. By utilizing canopy gaps, the results of this study compare favourably with the results obtained by Korpela et al. (2010), who used various intensity features for classification. The authors obtained an OA and KHAT of 90.8% and 0.84, respectively. Yu et al. (2014) obtained an OA of 62.1% using discrete LiDAR features for species classification. However, using a combination of full-waveform and discrete LiDAR features for classification, the authors obtained an OA of 73.4%. In this study using discrete LiDAR and a combination of intensity and texture features derived from canopy gaps obtained an OA of 90.91% for classification. The method used in this study yielded significant improvement on the results obtained by Cao et al. (2016). The authors obtained an OA of 75.8% and KHAT of 0.68 discriminating Masson pine, Chinese fir, Slash pines, Sawtooth oak, Sweet gum, and Chinese holly using full-waveform LiDAR features.

In this study, utilizing a combination of intensity and texture features and both forest canopy and canopy gaps obtained improved results compared with only utilizing intensity and texture information contained within canopy gaps. Intensity and texture information extracted from both forest canopy and canopy gaps obtained a OOB error of 3.66 and test accuracy of 94.74% (KHAT = 0.88) for classifying *E. grandis* and *E. dunnii*. These results indicate an improvement on the results obtained by Korpela et al. (2010) (OA = 90.8% and KHAT = 0.84). Furthermore, the results using a combination of forest canopy and canopy gaps show significant improvement on the results obtained by Cao et al. (2016) (OA = 75.8% and KHAT = 0.68).

The findings of this study suggest that using both intensity and texture features derived from one discrete LiDAR dataset has the capability to obtain accurate species classification results. Specifically, using intensity and texture information extracted from both forest canopy and canopy gaps.

4.4.2 An operational framework for species discrimination

Forest species classification using passive sensors has been well documented throughout literature. Although very high spatial resolution passive sensors have the ability to yield improved species classification results compared with low-medium spatial resolution passive sensors, the similarity of species spectral information is still problematic (Lucas et al. 2008; Ke et al. 2010; Adelabu and Dube 2015; Cho et al. 2015; Hill et al. 2010; Korpela et al. 2010). LiDAR overcomes this setback by providing accurate three-dimensional tree canopy information shown to yield promising species classification results (Kim et al. 2009b; Vauhkonen et al. 2009; Kim et al. 2011a).

The RF classifier has been documented to yield accurate forest species discrimination results and

often outperform other ensemble classifiers. LiDAR, in combination with RF, is particularly useful for forest species classification as shown in literature (Korpela et al. 2010; Yu et al. 2014; Adelabu and Dube 2015; Cao et al. 2016).

This study has showed that using the RF classifier and LiDAR intensity and texture information contained within forest canopy and canopy gaps can accurately discriminate forest species. An inherent limitation of this study is that canopy gaps were most prevalent within *E. grandis* and *E. dunnii*. The omission of a large quantity of canopy gaps within *E. smithii* limited the study to a binary classification. Additionally, the number of canopy gaps was also limited in compartment C8, thereby resulting in a smaller number of canopy gap polygons to use for information extraction for all compartments. Future studies should, therefore, investigate species discrimination within commercial plantations with an increased occurrence of canopy gaps, if possible. The commission of forest canopy polygons increased the total number of polygons utilized for LiDAR intensity and texture feature extraction and subsequently yielded improved *E. grandis* and *E. dunnii* classification results. However, the methodology presented may be improved by the addition of more features (such as spectral) and potentially yield improved species classification results.

The results of this study obtained accurate species classification results, particularly when using a combination of intensity and texture information within forest canopy and canopy gaps. Therefore, this methodology can potentially be operationalised within commercial forestry (i.e. to develop growth and yield models or to produce woodpuld and paper products) as well as within conservation sectors (i.e. for effective management of forest communities).

4.5 CONCLUSION

The overarching aim of this study was to discriminate *E. grandis* and *E. dunnii* within a commercial plantation using RF and LiDAR-derived intensity and texture features. Specifically, LiDAR intensity and texture features were extracted from canopy gaps for species discrimination. Additionally, the utility of information contained within both forest canopy and canopy gaps for species discrimination was also assessed. To the best of the knowledge of the author, the methodology used in this study is probably the first of its kind within a South African context and presents promising results, specifically when using a combination of the RF classifier and LiDAR intensity and texture information extracted from both forest canopy and canopy gaps. Therefore, this methodology is viable for operationalization in a commercial plantation.

CHAPTER 5: SYNTHESIS AND CONCLUSION

In this chapter, the overarching aim and objectives will be revisited and the scientific merits of the methodology will also be discussed. Furthermore, strengths, weaknesses and limitations of the methodology will be presented, followed by assumptions made and gaps identified in the study. The applicability of the techniques to other domains was investigated as well as the operational potential of the developed framework. Subsequently, recommendations for future research will be made, followed by a brief overview of data availability and accessibility. This chapter will be concluded with final concluding remarks.

5.1 THE AIM, OBJECTIVES, AND SCIENTIFIC MERITS THE RESEARCH

The overarching aim of the research was to investigate the utility of LiDAR for modelling forest canopy gaps and use the delineated canopy gaps for species modelling within a commercial plantation. The research entailed two main objectives addressed in Chapter 3 and Chapter 4, respectively. The first objective (chapter 3) was to detect and delineate forest canopy gaps using a canopy height model (CHM) and intensity raster within an object-based image analysis (OBIA) environment. Additionally, spatial analysis of canopy gaps was undertaken using Getis-Ord Gi* and FRAGSTATS. The second objective (chapter 4) was to discriminate *Eucalyptus grandis* and *Eucalyptus dunnii* using a random forest (RF) classifier and LiDAR-derived intensity and texture features. The utility of using canopy gaps and forest canopy for species classification was also assessed.

The methodology adopted for objectives 1 and 2 of this study yielded promising results. The second objective of this research also yielded promising species classification results using information contained within canopy gaps and using a combination of forest canopy and canopy gaps. The majority of literature utilized forest canopy information for species classification. Therefore, the work presented in Chapter 4 also presents novelty, particularly within South Africa. Additionally, the developed framework displayed robustness within a forestry plantation and the efficiency of the results may be of interest to forest managers and fellow researchers.

5.2 LIDAR FOR CANOPY GAP DELINEATION AND DETECTION AND SPECIES DISCRIMINATION USING DELINEATED CANOPY GAPS

In Chapter 3, the importance of canopy gaps was discussed, and various ways in which previous studies have undertaken canopy gap modelling were also presented. A novel approach using a LiDAR-derived CHM, an intensity raster, and a combination of the CHM and intensity raster were

also used to model canopy gaps within a *E. grandis* plantation. Using OBIA, canopy gaps were differentiated from forest canopy using multiresolution segmentation (MRS) and a rule-based classification. Furthermore, canopy gaps were spatially analysed using Getis-Ord Gi* and FRAGSTATS.

To detect and differentiate canopy gaps from surrounding forest canopy, MRS was undertaken using a scale factor of 5. Thereafter, a rule-based approach was utilized to model canopy gaps. The classification model was assessed using two methods i.e. thematic accuracy and comparative area-based assessment. In addition to undertaking canopy gap classification using the CHM and intensity raster independently, a combined dataset (i.e. CHM and intensity) was also evaluated. To assess the robustness of the methodology, this approach was tested on a second *E. grandis* compartment of a similar age.

The second objective was to undertake spatial analysis of canopy gaps. This analysis was undertaken on a block level using Getis-Ord Gi* and FRAGSTATS. Specifically, Getis-Ord Gi* was used to assess the spatial clustering of canopy gaps within block E, block F, as well as a combination of block E and F. Additionally, spatial characterisation of canopy gaps was assessed using four FRAGSTATS metrics. These include parentage of landscape (PLAND), patch density (PD), landscape shape index (LSI) and shape index.

In Chapter 4, various intensity and texture features were derived to discriminate *Eucalyptus grandis* and *Eucalyptus dunnii* in the commercial plantation. Firstly, various LiDAR intensity and texture features were generated. Subsequently, species were discriminated using a RF algorithm and LiDAR-derived intensity and texture information within canopy gaps. Additionally, the RF algorithm and LiDAR-derived intensity and texture information were extracted from both forest canopy and canopy gaps. The usefulness of LiDAR intensity and texture features for species discrimination was assessed independently as well as utilizing a combination of intensity and texture for classification.

Four compartments (two *E. grandis* compartments and two *E. dunnii* compartments) of varying ages were included in the analysis. Different compartment combinations were utilized to evaluate the influence of varying species ages on the output. Classification results were assessed using OOB error and test accuracies i.e. overall accuracy and KHAT.

The results obtained in the first objective, particularly using a combination of a LiDAR CHM and intensity raster, acquired promising results for canopy gap modelling. The second objective also showed promising results using LiDAR intensity and texture information within canopy gaps.

However, using a combination of canopy gaps and forest canopy yielded improved species discrimination results. The results of the methodology show that using LiDAR data is an efficient approach to model canopy gaps and to undertake species discrimination using delineated canopy gaps.

5.3 STRENGTHS, WEAKNESSES, AND LIMITATIONS OF TECHNIQUES

In Chapter 3, using the LiDAR-derived CHM and intensity raster independently for canopy gap classification obtained sufficient accuracies. Limitations of this methodology include the restricted amount of canopy gaps within the study area. Furthermore, utilizing the intensity raster for delineating canopy gaps often confuses forest canopy with canopy gaps in cases where canopy gaps were vegetated. Similarly, the presence of small waterbodies within canopy gaps was detected as forest canopy by the classifier using the intensity raster. However, utilizing a combination of the CHM and intensity raster overcame these limitations. By utilizing the combined dataset, promising canopy gap mapping results were yielded.

The term, canopy gap is a key concept within both components of this research i.e. Chapters 3 and 4. Therefore, similar to Chapter 3, the quantity of canopy gaps was also a limitation within Chapter 4. Most canopy gaps occurred within *E. grandis* and *E. dunnii* compartments, which constrained the analysis to a binary classification. Furthermore, compartment C8 contained the least amount of canopy gaps, which resulted in 30 canopy gaps to be utilized for all compartments for consistency. Despite these limitations, the results of the methodology proved to be promising when using canopy gaps for species discrimination. The addition of forest canopy polygons resulted in an increased sample for species discrimination using both forest canopy and canopy gaps. This larger sample obtained improved classification results compared with only using canopy gaps for species discrimination.

5.4 ASSUMPTIONS MADE AND GAPS IN THE STUDY

In this study, it was assumed that all trees within a compartment have the same height and similar textures. Additionally, it was assumed that the distribution of LiDAR points was uniform across the study area.

Both components of this research focused on using LiDAR data. A useful addition to this research would be aerial or multispectral imagery. For Chapter 3, the LiDAR-derived CHM and intensity raster could be compared with canopy gap delineation using multispectral data. The additional dataset could also be combined with the CHM and intensity raster to yield a combined dataset containing elevation, intensity and spectral information. The additional spectral information would

be beneficial to discriminate forest canopy from canopy gaps, particularly in cases with waterbodies.

Similarly, for Chapter 4, the additional aerial or multispectral imagery would be beneficial. The LiDAR-based approach used could also be evaluated against a spectral species discrimination approach in the plantation. Using the spectral information combined with the existing intensity and texture features could be combined to form a larger dataset for species classification. A feature selection algorithm could also be undertaken using this larger dataset to identify a subset of features to best discriminate forest species.

5.5 APPLICATION OF TECHNIQUES TO OTHER DOMAINS

The methodology presented in this research was applied to a commercial plantation. The methodology could potentially be applied within natural forests to detect and delineated canopy gaps. Additionally, this research can be employed in various other domains. Some of these include, detection of forest roads, where forest CHM information could potentially detect height differences between forest canopy and neighbouring roads. Furthermore, the results of Chapter 3 suggest that tree height assessment and tree crown delineation could also potentially be undertaken using an intensity raster and CHM.

The results of chapter 4, suggest that the techniques employed could potentially be applied to other domains such as crop type identification and differentiation, and urban land use identification. For example, the results of Chapter 4 suggest that using a combination of RF and LiDAR intensity and texture features could potentially differentiate different crop types, where the intra tree crown analysis of Li et al (2013a) may be applied to investigate intra crop type gaps. Additionally, urban classes such as buildings and open spaces could potentially be discriminated using intensity features, where varying intensity values between buildings and open spaces occur (Yan et al. 2015).

5.6 OPERATIONAL POTENTIAL OF DEVELOPED FRAMEWORK

Literature has reported the importance of canopy gap modelling and species discrimination within both conservation and commercial sectors, see for example Schliemann and Bockheim (2011), Muscolo et al. (2014), Wilson et al. (2012), and Immitzer et al. (2012). Varying accuracies for both canopy gap and species mapping have also been found using LiDAR data. The work presented in both components of this research showed novelty and robustness within a South African commercial plantation. The results achieved with the methodology often outperformed similar studies. Therefore, the methodology presented would be viable to be operationalised within

commercial forestry to assist in forest disturbance detection and forest species discrimination.

5.7 RECOMMENDATIONS FOR FUTURE RESEARCH, DATA AVAILABILITY AND ACCESSIBILITY

Future studies should investigate canopy gap modelling using a combination of multispectral imagery and LiDAR data within South African plantations. The addition of spectral information would be valuable in discriminating canopy gaps from surrounding forests, whereas LiDAR data would assist in eliminating the effect of shadows seen on multispectral imagery. The presence of a waterbody could also be mapped and masked using the presented approach.

Smaller canopy gaps in this study (i.e. less than 60m^2) were excluded in the analysis. Future studies should consider using these smaller studies. However, a clear distinction between inter-tree crown gaps and canopy gaps should be defined. Canopy gaps can also be further differentiated into naturally and harvest created canopy gaps and included in the classification.

Future studies should also investigate the addition of spectral information to the framework presented in Chapter 4. Additional features have the capability to obtain improved results compared with the results of Chapter 4. However, when using such a large dataset, the "curse of dimensionality" should be investigated.

The availability of remote sensing data is important for continuous and improved research, particularly for forestry research. Remote sensing data are becoming increasingly more available. Remote sensing data, such as SPOT and Landsat imagery can be requested from online data vendors at affordable prices, and some data vendors provide data freely. The price of LiDAR data is also decreasing.

5.8 CONCLUSIONS

Canopy gaps have conservation importance as these provide an opportunity for enhanced canopy growth by providing increased light availability, soil moisture, and nutrient availability (Negrón-Juárez et al. 2011; Muscolo et al. 2014). Within commercial plantations, canopy gaps could indicate where plant disease may occur within forest canopy. Forest managers would, therefore, benefit from accurate and timely canopy gap delineating techniques.

Accurate forest species classification is imperative for precise timber volume estimation for commercial forestry (Dalponte et al. 2008; Ko et al. 2013). In South Africa, *Eucalyptus* species are utilized for pulp and timber production (DAFF 2012). Accurate classification of these species

would be beneficial to the South African commercial forestry sector.

This study has shown that by using LiDAR data, canopy gaps can be accurately detected and delineate. Furthermore, using LiDAR intensity and texture features and the delineated canopy gaps and forest canopy information also yields accurate species classification results. Therefore, the objectives set out in this research were achieved.

CHAPTER 6: REFERENCES

Abdollahnejad A, Panagiotidis D, Joybari S & Surový P 2017. Prediction of Dominant Forest Tree Species Using QuickBird and Environmental Data. *Forests* 8, 2.

Adelabu S & Dube T 2015. Employing ground and satellite-based QuickBird data and random forest to discriminate five tree species in a Southern African Woodland. *Geocarto International* 30, 4: 457-471.

Aldred A & Hall J 1975. Application of large-scale photography to a forest inventory. *The Forestry Chronicle* 51: 9-15.

Andrew M, Ruthrof K, Matusick G & Hardy G 2016. Spatial configuration of drought disturbance and forest gap creation across environmental gradients. *PLoS ONE* 11, 6: 1-18.

Arenas-Castro S, Yulien Y, Jiménez-Muñoz J, Sobrino J, Fernández-Haeger J & Jordano-Bardudo D 2013. Mapping wild pear trees (*Pyrus bourgaeana*) in Mediterranean forest using high-resolution QuickBird satellite imagery. *International Journal of Remote Sensing* 34, 9-10: 3376-3396.

Baatz M & Schäpe A 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In Strobl J & Blaschke T (eds) Angewandte Geographische Informationsverarbeitung XII. Heidelberg, 12-23.

Baghdadi N, Cerdan O, Zribi M, Auzet V, Darboux F, Hajj M & Kheir R 2008. Operational performance of current synthetic aperture radar sensors in mapping soil surface characteristics in agriculture environments: application to hydrological and erosion modelling. *Hydrological Processing* 22: 9-20.

Barbati A, Chirici G, Corona P, Montaghi A & Travaglini D 2009. Area-based assessment of forest standing volume by field measurements and airborne laser scanner data. *International Journal of Remote Sensing* 30, 19: 5177-5194.

Barilotti A, Crosilla F & Sepic F 2009. Curvature analysis of LiDAR data for single tree species classification in alpine latitude forests. In Bretar F, Pierrot-Deseilligny M & Vosselman G (eds) *Laser scanning.* Proceedings of the IAPRS held 1-2 September 2009. Paris, France: International Archives of Photogrammetry and Remote Sensing.

Belgiu M & Drăguţ L 2014. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 96: 67-75.

Belgiu M & Drăguţ L 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114: 24-31.

Blaschke T 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65, 1: 2-16.

Blaschke T, Hay G, Kelly M, Lang S, Hofmann P, Addink E, Feitosa R, van der Meer F, van der Werff H, van Coillie F & Tiede D 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing* 87: 180-191.

Bonnet S, Gaulton R, Lehaire F & Lejeune P 2015. Canopy gap mapping from airborne laser scanning: An assessment of the positional and geometrical accuracy. *Remote Sensing* 7, 9: 11267-11294.

Bradley B & Fleishman E 2008. Can remote sensing of land cover improve species distribution modelling? *Journal of Biogeography* 35: 1158-1159.

Brandtberg T 2007. Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *ISPRS Journal of Photogrammetry & Remote Sensing* 61: 325-340.

Bredenkamp B & Upfold S 2012. *South African Forestry Handbook*. 5th ed. Menlo Park: South African Institue for Forestry.

Breiman L 2001. Random Forests. *Machine Learning* 45: 5-32.

Bright B, Hudak A, McGaughey R, Andersen H & Negrón J 2013. Predicting live and dead tree basal area of bark beetle affected forests from discrete-return lidar. *Canadian Journal of Remote Sensing* 39, S1: S99-S111.

Brokaw N 1982. The Definition of Treefall Gap and Its Effect on Measures of Forest Dynamics. *Biotropica* 14, 2: 158-160.

Brovelli M, Cannata M & Longani U 2004. LIDAR Data Filtering and DTM Interpolation Within GRASS. *Transaction in GIS* 8, 2: 155-174.

Cao L, Coops N, Innes J, Dai J, Ruan H & She G 2016. Tree species classification in subtropical forests using small-footprint full-waveform LiDAR data. *International Journal of Applied Earth Observation and Geoinformation* 49: 39-51.

Champion N, Matikainen L, Rottensteiner F, Liang X & Hyyppä J 2008. A test of 2D building change detection methods: Comparison, evaluation and perspectives. Proceedings of the ISPRS held 3-11 July 2008. Beijing, China: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences.

Chehata N, Guo L & Mallet C 2009. Airborne lidar feature selection for urban classification using random forests. In Bretar F, Pierrot-Deseilligny M & Vosselman G (eds) *Laser scanning.* Proceedings of the IAPRS held 1-2 September 2009. Paris, France: International Archives of Photogrammetry and Remote Sensing.

Chen W, Li X, Wang Y, Chen G & Liu S 2014. Forested landslide detection using LiDAR data and the random forest algorithm: A case study of the Three Gorges, China. *Remote Sensing of Environment* 152: 291-301.

Cho M, Malahlela O & Ramoelo A 2015. Assessing the utility WorldView-2 imagery for tree species mapping in South African subtropical humid forest and the conservation implications: Dukuduku forest patch as case study. *International Journal of Applied Earth Observation and Geoinformation* 38: 349-357.

Cho M, Mathieu R, Asner G, Naidoo L, van Aardt J, Ramoelo A, Debba P, Wessels K, Main R, Smit I & Erasmus B 2012. Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sensing of Environment* 125: 214-226.

Chuvieco E & Huete A 2010. *Fundamentals of satellite remote sensing*. Boca Raton: Taylor & Francis Group.

Congalton R & Green K 2009. *Assessing the accuracy of remotely sensed data: principles and practices*. 2nd ed. Boca Raton: CRC Press.

Corona P 2010. Integration of forest mapping and inventory to support forest management. *iForest- Biogeosciences and Forestry* 3: 59-64.

d'Oliveira M, Reutebuch S, McGaughey R & Andersen H 2012. Estimating forest biomass and identifying low-intensity logging areas using airborne scanning lidar in Antimary State Forest, Acre State, Western Brazilian Amazon. *Remote Sensing of Environment* 124: 479-491.

DAFF 2012. *Report on commercial timber resources and primary roundwood processing in South Africa 2011/2012*. Pretoria: Department of Agriculture, Forestry and Fisheries.

Dalponte M, Bruzzone L & Gianelle D 2008. Fusion of Hyperspectral and LIDAR Remote Sensing Data for Classification of Complex Forest Areas. *IEEE Transactions on Geoscience and Remote Sensing* 46, 5: 1416-1427.

Dalponte M, Bruzzone L & Gianelle D 2012. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sensing of Environment* 123: 258-270.

de Römer A, Kneeshaw D & Bergeron Y 2007. Small gap dynamics in the southern boreal forest of eastern Canada: Do canopy gaps influence stand development? *Journal of Vegetation Science* 18: 815-826.

Definiens 2007. *Definiens Developer 7- User Guide*. München: Definiens AG.

Dian Y, Li Z & Pang Y 2015. Spectral and Texture Features Combined for Forest Tree species Classification with Airborne Hyperspectral Imagery. *Journal of Indian Society of Remote Sensing* 43, 1: 101-107.

Dietze M & Clark J 2008. Changing the gap dynamics paradigm: Vegetative regeneration control on forest response to disturbance. *Ecological Monographs* 78, 3: 331-347.

Donoghue D, Watt P, Cox N & Wilson J 2007. Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sensing of Environment* 110: 509-522.

Drăguţ L, Tiede D & Levick S 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science* 24, 6: 859-871.

Einzmann K, Immitzer M, Böck S, Bauer O, Schmitt A & Atzberger C 2017. Windthrow detection in European forests with very high-resolution optical data. *Forests* 8, 1: 1-26.

Erdle K, Mistele B & Schmidhalter U 2011. Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. *Fields Crops Research* 124: 74-84.

Espírito-Santo F, Keller M, Linder E, Oliveira Junior R, Pereira C & Oliveira C 2014. Gap formation and carbon cycling in the Brazilian amazon: Measurement using high-resolution optical remote sensing and studies in large forest plots. *Plant Ecology Diversity* 7, 1–2: 305-318.

ESRI 2015. ArcGIS Pro [online]. Available from: http://pro.arcgis.com [Accessed 18 March 2017].

Evans J, Hudak A, Faux R & Smith A 2009. Discrete Return Lidar in Natural Resources: Recommendations for Project Planning, Data Processing, and Deliverables. *Remote Sensing* 1: 776-794.

Everitt J, Yang C, Sriharan S & Judd F 2008. Using High Resolution Satellite Imagery to Map Black Mangrove on the Texas Gulf Coast. *Journal of Coastal Research* 24, 6: 1582-1586.

Falkowski M, Evans J, Martinuzzi S, Gessler P & Hudak A 2009. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Remote Sensing of Environment* 113: 946-956.

Fitzgerald G 2010. Characterizing vegetation indices derived from active and passive sensors. *International Journal of Remote Sensing* 31, 16: 4335-4348.

Fox T, Knutson M & Hines R 2000. Mapping forest canopy gaps using air-photo interpretation and ground surveys. *Wildlife Society Bulletin* 28, 4: 882-889.

Frolking S, Palace M, Clark D, Chambers J, Shugart H, Hurtt G 2009. Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. *Journal of Geophysical Research Biogeosciences* 114, 3.

FUSION 2016. FUSION [online]. Seattle: University of Washington. Available from: http://forsys.cfr.washington.edu/fusion/fusionlatest.html [Accessed 3 June 2017].

Gao Y, Marpu P, Niemeyer I, Runfola D, Giner N, Hamill T & Pontius Jr. R 2011. Object-based classification with features extracted by a semi-automatic feature extraction algorithm – SEaTH. *Geocarto International* 26, 3: 211-226.

Garbarino M, Mondino E, Lingua E, Nagel T, Dukić V, Govedar Z & Motta R 2012. Gap disturbances and regeneration patterns in a Bosnian old-growth forest: a multispectral remote sensing and ground-based approach. *Annals of Forest Science* 69, 5: 617-625.

Gaulton R & Malthus T 2010. LiDAR mapping of canopy gaps in continuous cover forests: A comparison of canopy height model and point cloud based techniques. *International Journal of Remote Sensing* 31, 5: 1193-1211.

Gehler P & Nowozin S 2009. On feature combination for multiclass object classification. In ICCV. IEEE 12th International Conference on Computer Vision.

Getis A & Ord J 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24, 3: 189-206.

Gjertsen A 2007. Accuracy of forest mapping based on Landsat TM data and a kNN-based method. *Remote Sensing of Environment* 110: 420-430.

Gomes M & Maillard P 2013. Identification of urban tree crown in a tropical environment using WorldView-2 data: Problems and perspectives. Proceedings of SPIE held 23 September 2013. Dresden, Germany.

Gosh A, Fassnacht F, Joshi P & Koch B 2014. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *International Journal of Applied Earth Observation and Geoinformation* 26: 49-63.

Gray A, Spies T & Pabst R 2012. Canopy gaps affect long-term patterns of tree growth and mortality in mature and old-growth forests in the Pacific Northwest. *Forest Ecology Management* 281:111-120.

Haralick R, Shanmugam K & Dinstein I 1973. Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics* 3, 6: 610-621.

Heinzel J & Kock B 2011. Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation* 13: 152-160.

Hermosilla T, Ruiz L, Recio J & Estornell J 2011. Evaluation of automatic building detection approaches combining high resolution images and LiDAR data. *Remote Sensing* 3, 6: 1188-1210.

Hill R, Wilson A, George M & Hinsley S 2010. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Applied Vegetation Science* 13: 86-99.

Huang W, Li H & Lin G 2015. Classifying forest stands based on multi-scale structure features using Quickbird image. Proceedings of the 2nd IEEE ICSDM held 24-26 June 2015. Gdynia: International Conference on Spatial Data Mining and Geographical Knowledge.

Hunter G, Roux J, Wingfield B, Crous P & Wingfield M 2004. *Mycosphaerella* species causing leaf disease in South African *Eucalyptus* plantations. *Mycological Research* 108, 6: 672-681.

Hunter M, Keller M, Morton D, Cook B, Lefsky M, Ducey M, Saleska S, de Oliveira C & Schietti J 2015. Structural Dynamics of Tropical Moist Forest Gaps. *PLoS ONE* 10, 7: 1-19.

Hyyppä J, Hyyppä H, Leckie D, Gougeon F, Yu X & Maltamo M 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing* 29, 5: 1339-1366.

Immitzer M, Atzberger C & Koukal T 2012. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sensing* 4: 2661-2693.

Jensen D 2015. Identifying and Assessing the Yield Implications of Forest Canopy Gaps in Forest Management Using Full Feature LiDAR. Master's thesis. Alberta: University of Alberta, Department of Renewable Resources.

Kangas A & Maltamo M 2006. *Forest Inventory: Methodology and Applications (Managing Forest Ecosystems).* Dordrecht: Springer.

Kayitakire F, Hamel C & Defourny P 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment* 102: 390-401.

Ke Y, Quackenbush L & Im J 2010. Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. *Remote Sensing of Environment* 114: 1141-1154.

Kent R, Lindsell J, Laurin G, Valentini R & Coomes D 2015. Airborne LiDAR detects selectively logged tropical forest even in an advanced stage of recovery. *Remote Sensing* 7, 7: 8348-8367.

Key T, Warner T, McGraw J & Fajvan M 2001. A Comparison of Multispectral and Multitemporal Information in High Spatial Resolution Imagery for Classification of Individual Tree Species in a Temperate Hardwood Forest. *Remote Sensing of Environment* 75: 100-112.

Kim H & Sohn G 2010. 3D classification of power-line scene from airborne laser scanning data using random forests. In Paparoditis N, Pierrot-Deseilligny M, Mallet C & Tournaire O (eds) IAPRS XXXVIII held 1-3 September 2010. Saint-Mandé, France: International Archives of Photogrammetry and Remote Sensing.

Kim M, Madden M & Warner T 2009a. Forest Type Mapping using Object-specific Texture Measures from Multispectral Ikonos Imagery: Segmentation Quality and Image Classification Issues. *Photogrammetric Engineering & Remote Sensing* 75, 7: 819-829.

Kim M, Warner T, Madden M & Atkinson D 2011b. Multi-scale GEOBIA with very high spatial resolution digital aerial imagery: scale, texture and image objects. *International Journal of Remote Sensing* 32, 10: 2825-2850.

Kim S, Hinckley T & Briggs D 2011a. Classifying individual tree genera using stepwise cluster analysis based on height and intensity metrics derived from airborne laser scanner data. *Remote Sensing of Environment* 115: 3329-3342.

Kim S, McGaughey R, Andersen H & Schreuder G 2009b. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sensing of Environment* 113: 1575-1586.

Ko C, Sohn G & Remmel T 2013. Tree genera classification with geometric features from high-density airborne LiDAR. *Canadian Journal of Remote Sensing* 39, S1: S73-S85.

Korpela I, Ørka H, Maltamo M, Tokola T & Hyyppä J 2010. Tree species classification using airborne LiDAR – Effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica* 44, 2: 319-339.

Kotsiantis S 2011. Combining bagging, boosting, rotation forest and random subspace methods. *Artificial Intelligence Review* 35: 223-240.

Koukoulas S & Blackburn G 2004. Quantifying the spatial properties of forest canopy gaps using LiDAR imagery and GIS. *International Journal of Remote Sensing* 25, 15: 3049-3071.

Kucbel S, Jaloviar P, Saniga M, Vencurik J & Klimaš V 2010. Canopy gaps in an old-growth fir-beech forest remnant of Western Carpathians. *European Journal of Forest Research* 129, 3: 249-259.

Köhl M, Magnussen S & Marchetti M 2006. *Sampling Methods, Remote Sensing and GIS Multiresource Forest Inventory.* New York: Springer.

Lefsky M, Cohen W, Parker G & Harding D 2002. Lidar Remote Sensing for Ecosystem Studies. *BioScicence* 52, 1: 19-30.

Li J, Hu B & Noland T 2013a. Classification of tree species based on structural features derived from high density LiDAR data. *Agricultural and Forest Meteorology* 171-172: 104-114.

Li M, Im J & Beier C 2013b. Machine learning approaches for forest classification and change analysis using multi-temporal Landsat TM images over Huntington Wildlife Forest. *GIScience & Remote Sensing* 50, 4: 361-384.

Liaw A & Wiener M 2002. Classification and Regression by randomForest. *R News* 2, 3: 18-22.

Lillesand T, Kiefer R & Chipman J 2008. *Remote sensing and image interpretation.* 6th ed. Hoboken: John Wiley & Sons.

Lippitt C, Rogan J, Li Z, Eastman R & Jones T 2008. Mapping selective logging in mixed deciduous forest: A comparison of machine learning algorithms. *Photogrammetric Engineering & Remote Sensing* 74, 10: 1201-1211.

Liu D & Xia F 2010. Assessing object-based classification: advantages and limitations. *Remote Sensing Letter* 1, 4: 187-194.

Lombard L, Ismail R & Poona N 2017. Modelling forest canopy gaps using LiDAR-derived variables. *Geocarto International* DOI: 10.1080/10106049.2017.1377775.

Lucas R, Bunting P, Paterson M & Chrisholm L 2008. Classification of Australian forest communities using aerial photography, CASI and HyMap data. *Remote Sensing of Environment* 112: 2088-2103.

Mabvurira D & Miina J 2002. Individual-tree growth and mortality models for *Eucalyptus grandis* (Hill) Maiden plantations in Zimbabwe. *Forest Ecology and Management* 161, 1-3: 231-245.

Macfarlane D 2006. State of the environment report. KwaZulu-Natal: Sappi Forests.

Malahlela O, Cho M & Mutanga O 2014. Mapping canopy gaps in an indigenous subtropical coastal forest using high-resolution WorldView-2 data. *International Journal of Remote Sensing* 35, 17: 6397-6417.

Mallet C & Bretar F 2009. Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing* 64: 1-16.

Mallinis G, Koutsias N, Tsakiri-Strati M & Karteris M 2008. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. *ISPRS Journal of Photogrammetry and Remote Sensing* 63: 237-250.

Maltamo M, Næsset E & Vauhkonen J 2014. *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Dordrecht: Springer.

McGarigal K, Cushman S & Ene E 2012. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps [online]. Massachusetts: University of Massachusetts. Available from: http://www.umass.edu/landeco/research/fragstats/fragstats.html [Accessed 29 January 2017].

Meyer P, Staenz K & Itten K 1996. Semi-automated procedures for tree species identification in high spatial resolution data from digitized colour infrared-aerial photography. *ISPRS Journal of Photogrammetry & Remote Sensing* 51: 5-16.

Mhangara P & Odindi J 2013. Potential of texture-based classification in urban landscapes using multispectral aerial photos. *South African Journal of Science* 109, 3/4: 1-8.

Miao X, Heaton J, Zheng S, Charlet D & Liu H 2012. Applying tree-based ensemble algorithms to the classification of ecological zones multi-temporal multi-source remote-sensing data. *International Journal of Remote Sensing* 33, 6: 1823-1849.

Moffiet T, Mengersen K, Witte C, King R & Denham R 2005. Airborne laser scanning: Exploratory data analysis indicates potential variables for classification of individual trees or forest stands according to species. *ISPRS Journal of Photogrammetry & Remote Sensing* 59: 289-309.

Mulyani M & Jepson P 2017. Does the 'One Map Initiative' Represent a New Path for Forest Mapping in Indonesia? Assessing the Contribution of the REDD+ Initiative in Effecting Forest Governance Reform. *Forests* 8.

Muscolo A, Bagnato S, Sidari M & Mercurio R 2014. A review of the roles of forest canopy gaps. *Journal of Forestry Research* 25, 4: 725-736.

Navulur K 2007. *Multispectral image analysis using the object-oriented paradigm*. Boca Raton: CRC Press/Taylor & Francis.

Negrón-Juárez R, Chambers J, Marra D, Ribeiro G, Rifai S, Higuchi N & Roberts D 2011. Detection of subpixel treefall gaps with Landsat imagery in Central Amazon forests. *Remote Sensing of Environment* 115, 12: 3322-3328.

NOAA 2012. *"Lidar 101: An Introduction to Lidar Technology, Data, and Applications." Revised*. Charleston: National Oceanic and Atmospheric Administration Coastal Services Center.

Nussbaum S, Niemeyer I, Canty MJ. 2006. SEaTH - A new tool for automated feature extraction in the context of object-based image analysis. Proceedings of 1st International Conference on Object-based Image Analysis; Austria, Salzburg.

Odindi J, Mutangra O, Rouget M & Hlanguza N 2016. Mapping alien and indigenous vegetation in the KwaZulu-Natal Sandstone Sourveld using remotely sensed data. *Bothalia – African Biodiversity & Conservation* 46, 2: 1-9.

Olofsson K, Wallerman J, Holmgren J & Olsson H 2006. Tree species discrimination using Z/I DMC imagery and template matching of single trees. *Scandinavian Journal of Forest Research* 21, S7: 106-110.

Ørka H, Næsset E & Bollandsås O 2009. Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sensing of Environment* 113: 1163-1174.

Oza N & Tumer K 2008. Classifier ensembles: Select real-world applications. *Information Fusion* 9: 4-20.

Peerbhay K, Mutanga O & Ismail R 2013. Commercial tree species discrimination using airborne AISA Eagle hyperspectral imagery and partial least squares discriminant analysis (PLS-DA) in KwaZulu-Natal, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing* 79: 19-28.

Peerbhay K, Mutanga O & Ismail R 2014. Investigating the Capability of Few Strategically Placed Worldview-2 Multispectral Bands to Discriminate Forest Species in KwaZulu-Natal, South Africa. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing* 7, 1: 307-316.

Pekkarinen A, Reithmaier L & Stroble P 2009. Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data. *ISPRS Journal of Photogrammetry and Remote Sensing* 64: 171-183.

Poona N & Ismail R 2014. Using Boruta-selected spectroscopic wavebands for the asymptomatic detection of Fusarium circinatum stress. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7, 9: 3764-3772.

Popescu S, Zhao K, Neuenschwander A & Lin C 2011. Satellite lidar vs. small footprint airborne lidar: Comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level. *Remote Sensing of Environment* 115: 2786-2797.

Priestnall G, Jaafar J & Duncan A 2000. Extracting urban features from LiDAR digital surface models. *Computers, Environment and Urban Systems* 24: 65-78.

Puttonen E, Litkey P & Hyyppä J 2010. Individual Tree Species Classification by Illuminated-Shaded Area Separation. *Remote Sensing* 2: 19-35.

Qin Y, Xiao X, Wang J, Dong J, Ewing K, Hoagland B, Hough D, Fagin T, Zou Z, Geissler G, Xian G & Loveland T 2016. Mapping Annual Forest Cover in Sub-Humid and Semi-Arid Regions through Analysis of Landsat and PALSAR Imagery. *Remote Sensing* 8.

R Development Core Team 2017. R: A Language and Environment for Statistical Computing [online]. Vienna, Austria: The R Project for Statistical Computing. Available from: https://www.r-project.org/ [Accessed 8 June 2017].

Rahman M & Saha S 2008. Multi-resolution Segmentation for Object-based Classification and Accuracy Assessment of Land Use/Land Cover Classification using Remotely Sensed Data. *Journal of the Indian Society of Remote Sensing* 36: 189-201.

Räsänen A, Kuitunen M, Tomppo E & Lensu A 2014. Coupling high-resolution satellite imagery with ALS-based canopy height model and digital elevation model in object-based boreal forest habitat type classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 94: 169-182.

Reddy C, Manaswini G, Satish K, Singh S, Jha C & Dadhwal V 2016. Conservation priorities of forest ecosystems: Evaluation of deforestation and degradation hotspots using geospatial techniques. *Ecological Engineering* 91: 333-342.

Reitberger J, Krzystek P & Stilla U 2008. Analysis of the full waveform LIDAR data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing* 29, 5: 1407-1431.

Reutebuch S, Andersen H & McGaughey R 2005. Light Detection and Ranging (LIDAR): An Emerging Tool for Multiple Resource Inventory. *Journal of Forestry* 103: 286-292.

Richards J & Jia X 1999. *Remote sensing digital image analysis*. 3rd ed. Berlin: Springer.

Riley K, Grenfell I & Finney M 2016. Mapping forest vegetation for the western United States using modified random forests imputation of FIA forest plots. *Ecosphere* 7, 10.

Roberts J, Tesfamichael S, Gebrelasie M, van Aardt J & Ahmed F 2007. Forest structural assessment using remote sensing technologies: an overview of the current state of the art. *Southern Hemisphere Forestry Journal* 69, 3: 183-203.

Rodriguez-Galiano V, Ghimire B, Rogan J, Chica-Olmo M & Rigol-Sanchez J 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 93-104.

Sapkota I & Odén P 2009. Gap characteristics and their effects on regeneration, dominance and early growth of woody species. *Journal of Plant Ecology* 2, 1 : 21-9.

Sappi 2017. Sappi [online]. Available from: https://www.sappi.com/sappi-forests [Accessed 5 May 2017].

Säynäjoki R, Packalén P, Maltamo M, Vehams M & Eerikäinen K 2008. Detection of Aspens Using High Resolution Aerial Laser Scanning Data and Digital Aerial Images. *Sensors* 8: 5037-5054.

Schliemann S & Bockheim J 2011. Methods for studying treefall gaps: A review. *Forest Ecology and Management* 261, 7: 1143-1151.

Shang X & Chrisholm L 2014. Classification of Australian Native Forest Species Using Hyperspectral Remote Sensing and Machine-Learning Classification Algorithms. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7, 6: 2481-2489.

Shataee S, Kalbi S, Fallah A & Pelz D 2012. Forest attribute imputation using machine-learning methods and ASTER data: comparison of k-NN, SVR and random forest regression algorithms. *International Journal of Remote Sensing* 33, 19: 6254-6280.

Simonson W, Allen H & Coomes D 2012. Use of an Airborne Lidar System to Model Plant Species Composition and Diversity of Mediterranean Oak Forests. *Conservation Biology* 26, 5: 840-850.

Sing S 2013. Confused Between DEM, DTM and DSM ! [online]. Hyderabad: Jawaharlal Nehru Technological University. Available from: https://www.gisresources.com/confused-dem-dtm-dsm/ [Accessed 8 August 2017].

St-Onge B, Jumelet J, Cobello M & Véga C 2004. Measuring individual tree height using a combination of stereophotogrammetry and lidar. *Canadian Journal of Forest Research* 34: 2122-2130.

Suárez J, Smith S, Bull G, Malthus T, Donoghue D & Knox D 2005. The use of remote sensing techniques in operational forestry. *Quarterly Journal of Forestry* 99, 1: 31-42.

Swetnam T, Lynch A, Falk D, Yool S & Guertin D 2015. Discrimination disturbance from natural variation with LiDAR in semi-arid forests in the southwestern USA. *Ecosphere* 6: 1-22.

Tewari D 2001. Is commercial forestry sustainable in South Africa? The changing institutional and policy needs. *Forest Policy and Economics* 2: 333-353.

Tinkham W, Huang H, Smith A, Shrestha R, Falkowski M, Hudak A, Link T, Glenn N & Marks D 2011. A Comparison of Two Open Source LiDAR Surface Classification Algorithms. *Remote Sensing* 3: 638-649.

Trimble 2016. eCognition Trial Software [online]. Available from: http://www.ecognition.com/free-trial [Accessed 21 October 2016].

Vajari K, Jalilvand H, Espahbodi M & Moshki A 2012. Effect of canopy gap size and ecological factors on species diversity and beech seedlings in managed beech stands in Hyrcanian forests. *Journal of Forest Research* 23, 2: 217-222.

Van Coillie F, Verbeke L & De Wulf R 2007. Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium. *Remote Sensing of Environment* 110: 476-487.

Van Leeuwen & Nieuwenhuis M 2010. Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research* 129: 749-770.

Varo-Martínez M, Navarro-Cerrillo R, Hernández-Clemente R & Duque-Lazo J 2017. Semi-automated stand delineation in Mediterranean *Pinus sylvestris* plantations through segmentation of LiDAR data: The influence of pulse density. *International Journal of Applied Earth Observation and Geoinformation* 56: 54-64.

Vaughn N, Moskal L & Turnblom E 2012. Tree Species Detection Accuracies Using Discrete Point Lidar and Airborne Waveform Lidar. *Remote Sensing* 4: 377-403.

Vauhkonen J, Tokola T, Packalén & Maltamo M 2009. Identification of Scandinavian Commercial Species of Individual Trees from Airborne Laser Scanning Data Using Alpha Shape Metrics. *Forest Science* 55, 1: 37-47.

Vehmas M, Packalén P, Maltamo M & Eerikäinen K 2011. Using airborne laser scanning data for detecting canopy gaps and their understory type in mature boreal forest. *Annals of Forest Science* 68, 4: 825-833.

Vepakomma U, St-Onge B & Kneeshaw D 2008. Spatially explicit characterization of boreal forest gap dynamics using multi-temporal lidar data. *Remote Sensing of Environment* 112, 5: 2326-2340.

Vepakomma U, St-Onge B & Kneeshaw D 2011. Response of a boreal forest to canopy opening: Assessing vertical and lateral tree growth with multi-temporal lidar data. *Ecological Applications* 21, 1: 99-121.

Wagner W 2010. Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: Basic physical concepts. *ISPRS Journal of Photogrammetry and Remote Sensing* 65: 505-513.

Wagner W, Hollaus M, Briese C & Ducic V 2008. 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. *International Journal of Remote Sensing* 29, 5: 1433-1452.

Wang T, Zhang H, Lin H & Fang C 2016. Textural-Spectral Feature-Based Species Classification of Mangroves in Mai Po Nature Reserve from Wolrdview-3 Imagery. *Remote Sensing* 8.

Waser L, Fischer C, Wang Z & Ginzler C 2015. Wall-to-Wall Forest Mapping Based on Digital Surface Models from Image-Based Point Clouds and a NFI Forest Definition. *Forests* 6: 4510-4528.

Wilson B, Lister A & Riemann R 2012. A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *Forest Ecology and Management* 271: 182-198.

Wolter P & Townsend P 2011. Multi-sensor data fusion for estimating forest species composition and abundance in northern Minnesota. *Remote Sensing of Environment* 155: 671-691.

Wulder M, White J, Nelson R, Næsset E, Ørka H, Coops N, Hilker T, Bater C & Gobakken T 2012. Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment* 121: 196-2009.

Xie Y, Sha Z & Yu M 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology* 1, 1: 9-23.

Yan W, Shaker A & El-Ashmawy N 2015. Urban land cover classification using airborne LiDAR data: A review. *Remote Sensing of Environment* 158: 295-310.

Yang J, Jones T, Caspersen J & He Y 2015. Object-Based Canopy Gap Segmentation and Classification: Quantifying the Pros and Cons of Integrating Optical and LiDAR Data. *Remote Sensing* 7: 15917-15932.

Yang X, Rochdi N, Zhang J, Banting J, Rolfson D, King C, Staenz K, Patterson S & Purdy B 2014. Mapping tree species in a boreal forest area using RapidEye and LiDAR data. In IGARSS held 13 July 2014. International Geoscience and Remote Sensing Symposium.

Yu X, Litkey P, Hyyppä J, Holopainen M & Vastaranta M 2014. Assessment of Low Density Full-Waveform Airborne Laser Scanning for Individual Tree Detection and Tree Species Classification. *Forests* 5: 1011-1031.

Yunfei B, Guoping L, Chunxiang C, Xiaowen L, Hao Z, Qisheng H, Linyan B & Chaoyi C 2008. Classification of LIDAR point cloud and generation of DTM from LIDAR height and intensity data in forested area. Proceedings of IAPRS XXXVII. Beijing, China: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Zhang K 2008. Identification of gaps in mangrove forest with airborne LIDAR. *Remote Sensing of Environment* 112, 5: 2309-2325.

Zhao K, Popescu S, Meng X, Pang Y & Agca M 2011. Characterizing forest canopy structure with lidar composite metrics and machine learning. *Remote Sensing of Environment* 115: 1978-1996.

Zielewska-Büttner K, Adler P, Ehmann M & Braunisch V 2016. Automated detection of forest gaps in spruce dominated stands using canopy height models derived from stereo aerial imagery. *Remote Sensing* 8, 3.